

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/258,031A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2      Wrapped Aminos      The amino acid number/text at the end of each line "wrapped " down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3      Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4      Misaligned Amino Acid      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs  
Numbering      between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5      Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6      Variable Length      Sequence(s)      contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and  
indicate in the (ix) feature section that some may be missing.
- 7      PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid  
sequence(s)                     . Normally, PatentIn would automatically generate this section from the  
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section  
to the subsequent amino acid sequence.
- 8      Skipped Sequences      Sequence(s)      missing. If intentional, please use the following format for each skipped sequence:  
(OLD RULES)      (2) INFORMATION FOR SEQ ID NO:X:  
                    (i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
                    (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
                    This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9      Skipped Sequences      Sequence(s)      missing. If intentional, please use the following format for each skipped sequence.  
(NEW RULES)      <210> sequence id number  
                    <400> sequence id number  
                    000
- 10      Use of n's or Xaa's      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11      Use of <213>Organism      Sequence(s)              are missing this mandatory field or its response.  
(NEW RULES)
- 12      Use of <220>Feature      Sequence(s)      are missing the <220>Feature and associated headings.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13      PatentIn ver. 2.0 "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted  
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

R. Hutson

1652

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:33

INPUT SET: S33211.raw

This Raw Listing contains the General  
Information Section and those Sequences  
containing ERRORS.

Does Not Comply  
Corrected Diskette Needed

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: STUIVER, Maarten Hendrik  
CUSTERS, Jerome Humbertina Henricus Victor  
SELA-BURLAGE, Marianne Beatrix  
MELCHERS, Leo Sjoerd  
VAN DEVENTER-TROOST, Johanna Pieternella  
LAGEWEG, Wessel  
PONSTEIN, Anne Silene  
LAGEWEG, Wessel  
PONSTEIN, Anne Silene

(ii) TITLE OF INVENTION: ANTIFUNGAL PROTEINS, DNA CODING THEREFOR,  
AND HOSTS INCORPORATING SAME.

(iii) NUMBER OF SEQUENCES: 75

(iv) CORRESPONDENCE ADDRESS: (A) ADDRESSEE: LADAS & PARRY  
(B) STREET: 26 WEST 61 STREET  
(C) CITY: NEW YORK  
(D) NY STATE:  
(E) USA COUNTRY:  
(F) 10023 - 7604 ZIP:

insert a hard return - these subheadings  
appear  
UNDER  
main  
heading

(v) COMPUTER READABLE FORM: (A) MEDIUM TYPE: 3.25" Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: WINDOWS 95  
(D) SOFTWARE: WORDPERFECT 8

insert hard return

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/258,031  
(B) FILING DATE: 25-FEB-1999  
(C) CLASSIFICATION: 435

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/EP97/04923  
(B) FILING DATE: 04-SEP-1997

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: EP97200831.2  
(B) FILING DATE: 19-MAR-1997

(vii) PRIOR APPLICATION DATA:

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:33

INPUT SET: S33211.raw

46 (A) APPLICATION NUMBER: EP96202466.7  
47 (B) FILING DATE: 04-SEP-1996  
48

## ERRORED SEQUENCES FOLLOW:

150 (2) INFORMATION FOR SEQ ID NO: 5:

151

152 (i) SEQUENCE CHARACTERISTICS:

153 (A) LENGTH: 354 base pairs

154 (B) TYPE: nucleic acid

155 (C) STRANDEDNESS: double

156 (D) TOPOLOGY: linear

157

158 (ii) MOLECULE TYPE: DNA (genomic)

159

160 (iii) HYPOTHETICAL: NO

161

162 (iii) ANTI-SENSE: NO

163

164 (vi) ORIGINAL SOURCE:

165 (A) ORGANISM: Helianthus annuus

166 (B) STRAIN: cv. zebulon

167

168 (ix) FEATURE:

169 (A) NAME/KEY: CDS

170 (B) LOCATION: 1..354

171

172 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

173

174

175

176 GAT CCG TCT TTC CCG ATT ACT GGG GAG GTT TAC ACT CCC GGA AAC TCA 48

177 Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr Pro Gly Asn Ser

178 1 5 10 15

179

180 TCT TTT CCT ACC GTC TTG CAA AAC TAC ATC CGA AAC CTT CGG TTC AAT 96

181 Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn Leu Arg Phe Asn

182 20 25 30

183

184 GAA ACT ACC ACA CCA AAA CCC TTT TTA ATC ATC ACA GCC GAA CAT GTT 144

185 Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr Ala Glu His Val

186 35 40 45

187

188 TCC CAC ATT CAG GCA GCT GTG GTT TGT GGC AAA CAA AAC CGG TTG CTA 192

189 Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln Asn Arg Leu Leu

190 50 55 60

191

192 CTG AAA ACC AGA AGC GGT GGT CAT GAT TAT GAA GGT CTT TCC TAC CTT 240

193 Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly Leu Ser Tyr Leu

*(next page)*

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:34

INPUT SET: S33211.raw

```

194      65              70              75              80
195
196 ACA AAC ACA AAC CAA CCC TTC TTC ATT GTG GAC ATG TTC AAT TTA AGG      288
197 Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met Phe Asn Leu Arg
198              85              90              95
199
200 TCC ATA AAC GTA GAT ATC GAA CAA GAA ACC GCA TGG GTC CAA GCC GGC
201 336
202 Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp Val Gln Ala Gly
203              100              105              110
204
--> 205 GCC ACC CTC GGA GAA GTT
206 Ala Thr Leu Gly Glu Val
207              115
208
209

```

336 format error  
(see item 1 on Error summary sheet)

```

412 (2) INFORMATION FOR SEQ ID NO: 15:
413
414 (i) SEQUENCE CHARACTERISTICS:
--> 415 (A) LENGTH: 1784 base pairs
416 (B) TYPE: nucleic acid
417 (C) STRANDEDNESS: double
418 (D) TOPOLOGY: linear
419
420 (ii) MOLECULE TYPE: cDNA to mRNA
421
422 (iii) HYPOTHETICAL: NO
423
424 (iii) ANTI-SENSE: NO
425
426 (vi) ORIGINAL SOURCE:
427 (A) ORGANISM: Helianthus annuus
428 (B) STRAIN: Zebulon
429
430 (ix) FEATURE:
431 (A) NAME/KEY: CDS
432 (B) LOCATION: 21..1608
433
434 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

```

global error

```

435
436 ATATCACATC TTCTTTCAAC ATG CAA ACT TCC ATT CTT ACT CTC CTT CTT      50
437 Met Gln Thr Ser Ile Leu Thr Leu Leu Leu
438              1              5              10
439
440 CTC TTG CTC TCA ACC CAA TCT TCT GCA ACT TCC CGT TCC ATT ACA GAT      98
441 Leu Leu Leu Ser Thr Gln Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp
442              15              20              25
443
444 CGC TTC ATT CAA TGT TTA CAC GAC CGG GCC GAC CCT TCA TTT CCG ATA      146
445 Arg Phe Ile Gln Cys Leu His Asp Arg Ala Asp Pro Ser Phe Pro Ile
446              30              35              40
447

```

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:34

INPUT SET: S33211.raw

```

448 ACC GGA GAG GTT TAC ACT CCC GGA AAC TCA TCT TTT CCT ACC GTC TTG 194
449 Thr Gly Glu Val Tyr Thr Pro Gly Asn Ser Ser Phe Pro Thr Val Leu
450          45          50          55
451
452 CAA AAC TAC ATC CGA AAC CTT CGG TTC AAT GAA ACT ACC ACA CCA AAA 242
453 Gln Asn Tyr Ile Arg Asn Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys
454          60          65          70
455
456 CCC TTT TTA ATC ATC ACA GCC GAA CAT GTT TCC CAC ATT CAG GCA GCT 290
457 Pro Phe Leu Ile Ile Thr Ala Glu His Val Ser His Ile Gln Ala Ala
458          75          80          85          90
459
460 GTG GTT TGT GGC AAA CAA AAC CGG TTG CTA CTG AAA ACC AGA AGC GGT 338
461 338
462 Val Val Cys Gly Lys Gln Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly
463          95          100          105
--> 464 GGT CAT GAT TAT GAA GGT CTT TCC TAC CTT ACA AAC ACA AAC CAA CCC 386
465 Gly His Asp Tyr Glu Gly Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro
466          110          115          120
467
468 TTC TTC ATT GTG GAC ATG TTC AAT TTA AGG TCC ATA AAC GTA GAT ATC 434
469 Phe Phe Ile Val Asp Met Phe Asn Leu Arg Ser Ile Asn Val Asp Ile
470          125          130          135
471
472 GAA CAA GAA ACC GCA TGG GTC CAA GCC GGT GCG ACT CTT GGT GAA GTG
473 482
474 Glu Gln Glu Thr Ala Trp Val Gln Ala Gly Ala Thr Leu Gly Glu Val
475          140          145          150
476
477 TAC TAT CGA ATA GCG GAG AAA AGT AAC AAG CAT GGT TTT CCG GCA GGG
478 530
479 Tyr Tyr Arg Ile Ala Glu Lys Ser Asn Lys His Gly Phe Pro Ala Gly
480 155          160          165          170
481
--> 482 GTT TGT CCA ACG GTT GGC GTT GGT GGG CAT TTT AGT GGT GGT GGG TAT 578
483 Val Cys Pro Thr Val Gly Val Gly Gly His Phe Ser Gly Gly Gly Tyr
484          175          180          185
485
486 GGT AAT TTG ATG AGA AAA TAT GGT TTG TCG GTT GAT AAT ATT GTT GAT 626
487 Gly Asn Leu Met Arg Lys Tyr Gly Leu Ser Val Asp Asn Ile Val Asp
488          190          195          200
489
490 GCT CAA ATA ATA GAT GTG AAT GGC AAG CTT TTG GAT CGA AAG AGT ATG 674
491 Ala Gln Ile Ile Asp Val Asn Gly Lys Leu Leu Asp Arg Lys Ser Met
492          205          210          215
493
494 GGT GAG GAT TTG TTT TGG GCG ATC ACC GGC GGT GGT GGT GTT AGT TTT 722
495 Gly Glu Asp Leu Phe Trp Ala Ile Thr Gly Gly Gly Gly Val Ser Phe
496          220          225          230
497
498 GGT GTG GTT CTA GCC TAC AAA ATC AAA CTA GTT CGT GTT CCG GAG GTT 770
499 Gly Val Val Leu Ala Tyr Lys Ile Lys Leu Val Arg Val Pro Glu Val
500 235          240          245          250

```

338

same  
error

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:35

INPUT SET: S33211.raw

```

501
502 GTG ACC GTG TTT ACC ATT GAA AGA AGA GAG GAA CAA AAC CTC AGC ACC same
503 818
504 Val Thr Val Phe Thr Ile Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr
505 255 260 265
506
--> 507 ATC GCG GAA CGA TGG GTA CAA GTT GCT GAT AAG CTA GAT AGA GAT CTT 866
508 Ile Ala Glu Arg Trp Val Gln Val Ala Asp Lys Leu Asp Arg Asp Leu
509 270 275 280
510
511 TTC CTT CGA ATG ACC TTT AGT GTC ATA AAC GAT ACC AAC GGT GGA AAG 914
512 Phe Leu Arg Met Thr Phe Ser Val Ile Asn Asp Thr Asn Gly Gly Lys
513 285 290 295
514
515 ACA GTC CGT GCT ATC TTT CCA ACG TTG TAC CTT GGA AAC TCG AGG AAT 962
516 Thr Val Arg Ala Ile Phe Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn
517 300 305 310
518
519 CTT GTT ACA CTT TTG AAT AAA GAT TTC CCC GAG TTA GGG TTG CAA GAA 1010
520 Leu Val Thr Leu Leu Asn Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu
521 315 320 325 330
522
523 TCG GAT TGT ACT GAA ATG AGT TGG GTT GAG TCT GTG CTT TAC TAC ACG 1058
524 Ser Asp Cys Thr Glu Met Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr
525 335 340 345
526
527 GGC TTC CCC AGT GGT ACT CCA ACC ACG GCG CTC TTA AGC CGT ACT CCT 1106
528 Gly Phe Pro Ser Gly Thr Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro
529 350 355 360
530
531
532
533 CAA AGA CTC AAC CCA TTC AAG ATC AAA TCC GAT TAT GTG CAA AAT CCT 1154
534 Gln Arg Leu Asn Pro Phe Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro
535 365 370 375
536
537 ATT TCT AAA CGA CAG TTC GAG TTC ATC TTC GAA AGG CTG AAA GAA CTT 1202
538 Ile Ser Lys Arg Gln Phe Glu Phe Ile Phe Glu Arg Leu Lys Glu Leu
539 380 385 390
540
541 GAA AAC CAA ATG TTG GCT TTC AAC CCA TAT GGT GGT AGA ATG AGT GAA
542 1250
543 Glu Asn Gln Met Leu Ala Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu same
544 395 400 405 410
545
--> 546 ATA TCC GAA TTC GCA AAG CCT TTC CCA CAT AGA TCG GGT AAC ATA GCG 1298
547 Ile Ser Glu Phe Ala Lys Pro Phe Pro His Arg Ser Gly Asn Ile Ala
548 415 420 425
549
550 AAA ATT CAA TAC GAA GTA AAC TGG GAG GAT CTT AGC GAT GAA GCC GAA
551 1346
552 Lys Ile Gln Tyr Glu Val Asn Trp Glu Asp Leu Ser Asp Glu Ala Glu same
553 430 435 440

```

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:35

INPUT SET: S33211.raw

```

554
--> 555 AAT CGT TAC TTG AAT TTC ACA AGG CTG ATG TAT GAT TAC ATG ACC CCA 1394
556 Asn Arg Tyr Leu Asn Phe Thr Arg Leu Met Tyr Asp Tyr Met Thr Pro
557 445 450 455
558
559 TTT GTG TCG AAA AAC CCT AGA AAA GCA TTT TTG AAC TAT AGG GAT TTG 1442
560 Phe Val Ser Lys Asn Pro Arg Lys Ala Phe Leu Asn Tyr Arg Asp Leu
561 460 465 470
562
563 GAT ATT GGT ATC AAC AGC CAT GGC AGG AAT GCT TAT ACT GAA GGA ATG
564 1490
565 Asp Ile Gly Ile Asn Ser His Gly Arg Asn Ala Tyr Thr Glu Gly Met
566 475 480 485 490
567
568 GTT TAT GGG CAC AAG TAT TTC AAA GAG ACA AAT TAC AAG AGG CTA GTA
569 1538
570 Val Tyr Gly His Lys Tyr Phe Lys Glu Thr Asn Tyr Lys Arg Leu Val
571 495 500 505
572
573 AGT GTG AAG ACT AAA GTT GAT CCT GAC AAC TTC TTT AGG AAT GAG CAA
574 1586
575 Ser Val Lys Thr Lys Val Asp Pro Asp Asn Phe Phe Arg Asn Glu Gln
576 510 515 520
577
578 AGC ATC CCA ACT TTG TCA TCT T GAAGAACGTA CATATATAAA TAAATACCTT
579 1638
580 Ser Ile Pro Thr Leu Ser Ser
581 525
582
583 TGTCATGGT ATTTTCAGGG TGTTAAAGTG ATATTCAGAT ATTTATGATA
--> 584 GAATTTTGAC 1698
585
586 TTGTATTTTA TACAATCAAA ATTGTATGGT TCTCCGAATT TCTCTTTTTA
--> 587 ATTCTGAAAA 1758
588
589 ATACATATTA GTATTGTCAA AAAAAA 1784
590
591

```

*same*  
↓

```

743 (2) INFORMATION FOR SEQ ID NO: 19:
744
745 (i) SEQUENCE CHARACTERISTICS:
--> 746 (A) LENGTH: 1589 base pairs
747 (B) TYPE: nucleic acid
748 (C) STRANDEDNESS: double
749 (D) TOPOLOGY: linear
750
751 (ii) MOLECULE TYPE: cDNA to mRNA
752
753 (iii) HYPOTHETICAL: NO
754
755 (iii) ANTI-SENSE: NO
756

```

*PR-8-9*

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:35

INPUT SET: S33211.raw

```

757      (vi) ORIGINAL SOURCE:
758          (A) ORGANISM: Helianthus annuus
759          (B) STRAIN: Zebulon
760
761      (ix) FEATURE:
762          (A) NAME/KEY: CDS
763          (B) LOCATION: 1..1590
764
765      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
766
767      ATG GAG ACT TCC ATT CTT ACT CTC CTT CTT CTC TTG CTC TCA ACC CAA      48
768      Met Glu Thr Ser Ile Leu Thr Leu Leu Leu Leu Leu Ser Thr Gln
769          1              5              10              15
770
771
772
773      TCT TCT GCA ACT TCC CGT TCC ATT ACA GAT CGC TTC ATT CAA TGT TTA      96
774      Ser Ser Ala Thr Ser Arg Ser Ile Thr Asp Arg Phe Ile Gln Cys Leu
775          20              25              30
776
777      CAC GAC CGG GCC GAC CCT TCA TTT CCG ATA ACC GGA GAG GTT TAC ACT      144
778      His Asp Arg Ala Asp Pro Ser Phe Pro Ile Thr Gly Glu Val Tyr Thr
779          35              40              45
780
781      CCC GGA AAC TCA TCT TTT CCT ACC GTC TTG CAA AAC TAC ATC CGA AAC      192
782      Pro Gly Asn Ser Ser Phe Pro Thr Val Leu Gln Asn Tyr Ile Arg Asn
783          50              55              60
784
785      CTT CGG TTC AAT GAA ACT ACC ACA CCA AAA CCC TTT TTA ATC ATC ACA      240
786      Leu Arg Phe Asn Glu Thr Thr Thr Pro Lys Pro Phe Leu Ile Ile Thr
787          65              70              75              80
788
789      GCC GAA CAT GTT TCC CAC ATT CAG GCA GCT GTG GTT TGT GGC AAA CAA      288
790      Ala Glu His Val Ser His Ile Gln Ala Ala Val Val Cys Gly Lys Gln
791          85              90              95
792
793      AAC CGG TTG CTA CTG AAA ACC AGA AGC GGT GGT CAT GAT TAT GAA GGT      336
794      Asn Arg Leu Leu Leu Lys Thr Arg Ser Gly Gly His Asp Tyr Glu Gly
795          100             105             110
796
797      CTT TCC TAC CTT ACA AAC ACA AAC CAA CCC TTC TTC ATT GTG GAC ATG      384
798      Leu Ser Tyr Leu Thr Asn Thr Asn Gln Pro Phe Phe Ile Val Asp Met
799          115             120             125
800
801      TTC AAT TTA AGG TCC ATA AAC GTA GAT ATC GAA CAA GAA ACC GCA TGG      432
802      Phe Asn Leu Arg Ser Ile Asn Val Asp Ile Glu Gln Glu Thr Ala Trp
803          130             135             140
804
805      GTC CAA GCC GGT GCG ACT CTT GGT GAA GTG TAC TAT CGA ATA GCG GAG      480
806      Val Gln Ala Gly Ala Thr Leu Gly Glu Val Tyr Tyr Arg Ile Ala Glu
807          145             150             155             160
808
809      AAA AGT AAC AAG CAT GGT TTT CCG GCA GGG GTT TGT CCA ACG GTT GGC      528

```



# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:36

INPUT SET: S33211.raw

810	Lys Ser Asn Lys His Gly Phe Pro Ala Gly Val Cys Pro Thr Val Gly	
811		165 170 175
812		
813	GTT GGT GGG CAT TTT AGT GGT GGT GGG TAT GGT AAT TTG ATG AGA AAA	576
814	Val Gly Gly His Phe Ser Gly Gly Gly Tyr Gly Asn Leu Met Arg Lys	
815		180 185 190
816		
817	TAT GGT TTG TCG GTT GAT AAT ATT GTT GAT GCT CAA ATA ATA GAT GTG	624
818	Tyr Gly Leu Ser Val Asp Asn Ile Val Asp Ala Gln Ile Ile Asp Val	
819		195 200 205
820		
821	AAT GGC AAG CTT TTG GAT CGA AAG AGT ATG GGT GAG GAT TTG TTT TGG	672
822	Asn Gly Lys Leu Leu Asp Arg Lys Ser Met Gly Glu Asp Leu Phe Trp	
823		210 215 220
824		
825	GCG ATC ACC GGC GGT GGT GGT GTT AGT TTT GGT GTG GTT CTA GCC TAC	720
826	Ala Ile Thr Gly Gly Gly Gly Val Ser Phe Gly Val Val Leu Ala Tyr	
827		225 230 235 240
828		
829		
830		
831	AAA ATC AAA CTA GTT CGT GTT CCG GAG GTT GTG ACC GTG TTT ACC ATT	768
832	Lys Ile Lys Leu Val Arg Val Pro Glu Val Val Thr Val Phe Thr Ile	
833		245 250 255
834		
835	GAA AGA AGA GAG GAA CAA AAC CTC AGC ACC ATC GCG GAA CGA TGG GTA	
836	816	
837	Glu Arg Arg Glu Glu Gln Asn Leu Ser Thr Ile Ala Glu Arg Trp Val	
838		260 265 270
839		
--> 840	CAA GTT GCT GAT AAG CTA GAT AGA GAT CTT TTC CTT CGA ATG ACC TTT	864
841	Gln Val Ala Asp Lys Leu Asp Arg Asp Leu Phe Leu Arg Met Thr Phe	
842		275 280 285
843		
844	AGT GTC ATA AAC GAT ACC AAC GGT GGA AAG ACA GTC CGT GCT ATC TTT	912
845	Ser Val Ile Asn Asp Thr Asn Gly Gly Lys Thr Val Arg Ala Ile Phe	
846		290 295 300
847		
848	CCA ACG TTG TAC CTT GGA AAC TCG AGG AAT CTT GTT ACA CTT TTG AAT	960
849	Pro Thr Leu Tyr Leu Gly Asn Ser Arg Asn Leu Val Thr Leu Leu Asn	
850		305 310 315 320
851		
852	AAA GAT TTC CCC GAG TTA GGG TTG CAA GAA TCG GAT TGT ACT GAA ATG	
853	1008	
854	Lys Asp Phe Pro Glu Leu Gly Leu Gln Glu Ser Asp Cys Thr Glu Met	
855		325 330 335
856		
--> 857	AGT TGG GTT GAG TCT GTG CTT TAC TAC ACG GGC TTC CCC AGT GGT ACT	1056
858	Ser Trp Val Glu Ser Val Leu Tyr Tyr Thr Gly Phe Pro Ser Gly Thr	
859		340 345 350
860		
861	CCA ACC ACG GCG CTC TTA AGC CGT ACT CCT CAA AGA CTC AAC CCA TTC	1104
862	Pro Thr Thr Ala Leu Leu Ser Arg Thr Pro Gln Arg Leu Asn Pro Phe	

*same*

*same*

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

 DATE: 09/03/1999  
 TIME: 14:51:36

INPUT SET: S33211.raw

	863	355	360	365	
864					
865	AAG ATC AAA TCC GAT TAT GTG CAA AAT CCT ATT TCT AAA CGA CAG TTC	1152			
866	Lys Ile Lys Ser Asp Tyr Val Gln Asn Pro Ile Ser Lys Arg Gln Phe				
867	370 375 380				
868					
869	GAG TTC ATC TTC GAA AGG ATG AAA GAA CTT GAA AAC CAA ATG TTG GCG				
870	1200				
871	Glu Phe Ile Phe Glu Arg Met Lys Glu Leu Glu Asn Gln Met Leu Ala				
872	385 390 395 400				
873					
874	TTC AAC CCA TAT GGT GGT AGA ATG AGT GAA ATA TCC GAA TTC GCA AAG				
875	1248				
876	Phe Asn Pro Tyr Gly Gly Arg Met Ser Glu Ile Ser Glu Phe Ala Lys				
877	405 410 415				
878					
879	CCT TTC CCA CAT AGA TCG GGT AAC ATA GCG AAG ATT CAA TAC GAA GTA				
880	1296				
881	Pro Phe Pro His Arg Ser Gly Asn Ile Ala Lys Ile Gln Tyr Glu Val				
882	420 425 430				
883					
--> 884	AAC TGG GAG GAT CTT AGC GAT GAA GCC GAA AAT CGT TAC TTG AAT TTC	1344			
885	Asn Trp Glu Asp Leu Ser Asp Glu Ala Glu Asn Arg Tyr Leu Asn Phe				
886	435 440 445				
887					
888	ACA AGG CTG ATG TAT GAT TAC ATG ACT CCA TTT GTG TCG AAA AAC CCT	1392			
889	Thr Arg Leu Met Tyr Asp Tyr Met Thr Pro Phe Val Ser Lys Asn Pro				
890	450 455 460				
891					
892					
893					
894	AGA GAA GCA TTT TTG AAC TAT AGG GAT TTG GAT ATT GGT ATC AAC AGC	1440			
895	Arg Glu Ala Phe Leu Asn Tyr Arg Asp Leu Asp Ile Gly Ile Asn Ser				
896	465 470 475 480				
897					
898	CAT GGC AGG AAT GCT TAT ACT GAA GGA ATG GTT TAT GGG CAC AAA TAT				
899	1488				
900	His Gly Arg Asn Ala Tyr Thr Glu Gly Met Val Tyr Gly His Lys Tyr				
901	485 490 495				
902					
903	TTC AAA GAG ACA AAT TAC AAG AGG CTA GTA AGT GTG AAG ACT AAA GTT				
904	1536				
905	Phe Lys Glu Thr Asn Tyr Lys Arg Leu Val Ser Val Lys Thr Lys Val				
906	500 505 510				
907					
--> 908	GAT CCT GAC AAC TTC TTT AGG AAT GAG CAA AGC ATC CCA ACT TTG TCA	1584			
909	Asp Pro Asp Asn Phe Phe Arg Asn Glu Gln Ser Ile Pro Thr Leu Ser				
910	515 520 525				
911					
912	TCT TG	1589			
913	Ser				
914	530				
915					

INPUT SET: S33211.raw

916

---

1031 (2) INFORMATION FOR SEQ ID NO: 21:  
1032  
1033 (i) SEQUENCE CHARACTERISTICS:  
--> 1034 (A) LENGTH: 350 base pairs  
1035 (B) TYPE: nucleic acid  
1036 (C) STRANDEDNESS: double  
1037 (D) TOPOLOGY: linear  
1038  
1039 (ii) MOLECULE TYPE: cDNA to mRNA  
1040  
1041 (iii) HYPOTHETICAL: NO  
1042  
1043 (iii) ANTI-SENSE: NO  
1044  
1045 (vi) ORIGINAL SOURCE:  
1046 (A) ORGANISM: Arabidopsis thaliana  
1047 (B) STRAIN: ecotype Columbia  
1048 (ix) FEATURE:  
1049 (A) NAME/KEY: CDS  
1050 (B) LOCATION: 2..350  
1051  
1052 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
1053  
1054 GAGAAACTCG GAGACTTTCA CACAATGCCT AACCTCAAAC TCCGACCCCA  
--> 1055 AACATCCCAT 60  
1056  
1057 CTCCCCGCT ATCTTCTTCT CCGGAAATGG CTCCTACTCC TCCGTATTAC  
--> 1058 AAGCCAACAT 120  
1059  
1060 CCGTAACCTC CGCTTCAACA CCACCTCAAC TCCGAAACCC TTCCTCATAA  
--> 1061 TCGCCGCAAC 180  
1062  
1063 ACATGAATCC CATGTGCAAG CCGCGATTAC TTGCGGGAAA CGCCACAACC  
--> 1064 TTCAGATGAA 240  
1065  
1066 AATCAGAAGT GGAGGCCACG ACTACGATGG CTTGTCATAC GTTACATACT  
--> 1067 CTGGCAAACC 300  
1068  
1069 GTTCTTCGTC CTCGACATGT TTAACCTCCG TTCGGTGGAT GTCGACGTGG  
1070  
1071

350

---

1072 (2) INFORMATION FOR SEQ ID NO: 22:  
1073  
1074 (i) SEQUENCE CHARACTERISTICS:  
--> 1075 (A) LENGTH: 278 base pairs  
1076 (B) TYPE: nucleic acid  
1077 (C) STRANDEDNESS: double  
1078 (D) TOPOLOGY: linear  
1079  
1080 (ii) MOLECULE TYPE: cDNA to mRNA

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999

TIME: 14:51:37

INPUT SET: S33211.raw

1081  
1082 (iii) HYPOTHETICAL: NO  
1083  
1084 (iii) ANTI-SENSE: NO  
1085  
1086 (vi) ORIGINAL SOURCE:  
1087 (A) ORGANISM: Arabidopsis thaliana  
1088 (B) STRAIN: ecotype Columbia  
1089  
1090 (ix) FEATURE:  
1091 (A) NAME/KEY: CDS  
1092 (B) LOCATION: 2..278  
1093  
1094 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
1095  
1096 GGCATGGATC TCCGCCGGAG CGACTCTCGG AGAGGTTTAT TATCGGATTT  
--> 1097 GGGAGAAAAG 60  
1098  
1099 CAGAGTCCAT GGATTCCCCG CCGGAGTTTG ACCGACGGTT GGTGTTGGTG  
--> 1100 GGCATTTAAG 120  
1101  
1102 CGGCGGTGGT TACGGTAACA TGGTGAGGAA GTTTGGATTA TCTGTGGATT  
--> 1103 ACGTTGAGGA 180  
1104  
1105 TGCCAAGATC GTCGATGTAA ACNGTCGGGT TTTAGATCGG AAAGCAATGG  
--> 1106 GTGAGGATCT 240  
1107  
1108 GTTCTGGGCG ATTACCGGTG GAGGAGGAGG TAGCGTAC  
1109  
1110

278

---

1111 (2) INFORMATION FOR SEQ ID NO: 23:  
1112  
1113 (i) SEQUENCE CHARACTERISTICS:  
--> 1114 (A) LENGTH: 345 base pairs  
1115 (B) TYPE: nucleic acid  
1116 (C) STRANDEDNESS: double  
1117 (D) TOPOLOGY: linear  
1118  
1119 (ii) MOLECULE TYPE: cDNA  
1120  
1121 (iii) HYPOTHETICAL: NO  
1122  
1123 (iii) ANTI-SENSE: NO  
1124  
1125 (vi) ORIGINAL SOURCE:  
1126 (A) ORGANISM: Arabidopsis thaliana  
1127 (B) STRAIN: ecotype Columbia  
1128  
1129 (ix) FEATURE:  
1130 (A) NAME/KEY: CDS  
1131 (B) LOCATION: 2..345  
1132

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:37

INPUT SET: S33211.raw

1133 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
1134  
1135 TGGACATATT AGCGGAGGAG GATTCGGTAC AATAATGAGG AAATACGGTT  
--> 1136 TAGCGTCTGA 60  
1137  
1138 TAACGTTGTG GACGCACGTT TGATGGATGT AAATGGGAAA ACTCTTGACC  
--> 1139 GGAAAAACGAT 120  
1140  
1141 GGGAGAGGAT TTGTTTTGGG CGCTTAGAGG CGGTGGAGCT GCGAGTTTTG  
--> 1142 GCGTTGTCTT 180  
1143  
1144 GTCGTGGAAG GTTAAGCTTG CTAGGGTTCC TGAAAAGGTA ACTTGTTTCA  
--> 1145 TAAGTCAACA 240  
1146  
1147 TCCGATGGGA CCTAGCATGA ACAAGCTTGT TCATAGATGG CAATCCATAG  
--> 1148 GATCAAGANN 300  
1149  
1150 GCTAGACGAA GATTTATTCA TCAGAGTCAA TATTGACAAC AGTCT  
1151  
1152

345

---

1153 (2) INFORMATION FOR SEQ ID NO: 24:  
1154  
1155 (i) SEQUENCE CHARACTERISTICS:  
--> 1156 (A) LENGTH: 695 base pairs  
1157 (B) TYPE: nucleic acid  
1158 (C) STRANDEDNESS: double  
1159 (D) TOPOLOGY: linear  
1160  
1161 (ii) MOLECULE TYPE: cDNA to mRNA  
1162  
1163 (iii) HYPOTHETICAL: NO  
1164  
1165 (iii) ANTI-SENSE: NO  
1166  
1167 (vi) ORIGINAL SOURCE:  
1168 (A) ORGANISM: Arabidopsis thaliana  
1169 (B) STRAIN: ecotype Columbia  
1170  
1171 (ix) FEATURE:  
1172 (A) NAME/KEY: CDS  
1173 (B) LOCATION: 1..695  
1174  
1175 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
1176  
1177 GTTCGTTAAA ACCTATCCTN NANGGGCNAA AGNATATCAA AGNTTGNTTA  
--> 1178 NGNAACCCAA 60  
1179  
1180 NATTTCTGAA CTGGCCNCCT TCGGTGGTAT ATGNCNAAAN CCCTTGAATC  
--> 1181 TGCGNANCCN 120  
1182  
1183 ATTCCGCATA GAAACGGAAC CCTCTTCAAG ATTCTCTATT TACNCGAACT  
--> 1184 GNCTAGANNG 180

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:37

INPUT SET: S33211.raw

1185  
1186 AATGACAAGA CATCGAGTAG NAAAATCAAC TGGATCAAAG AGATATACAA  
--> 1187 TTACATGGCG 240  
1188  
1189 CCTTATGTCT CAAGCAATCC AAGACAAGCA TATGTGAACT ACAGAGATCT  
--> 1190 AGACTTCGGA 300  
1191  
1192 CAGAACAAGA ACAACGCAAA GGTAACTTC ATTGAAGCTA AAATCTGGGG  
--> 1193 ACCTAAGTAC 360  
1194  
1195 TTCAAAGGCA ATTTTGACAG ATTGGTGAAG ATTAAAACCA AGGTTGATCC  
--> 1196 AGAGAACTTC 420  
1197  
1198 TTCAGGCACG AGCAGAGTAT CCCACCTATG CCCTACTAGA AGCTAGGTTT  
--> 1199 ATGAAACCAA 480  
1200  
1201 TAACATTATC AAAAATAAGR ATAAATGRTA ATTGTATACA ACATGATTCG  
--> 1202 KCTTTCTTTA 540  
1203  
1204 TTTCAGACAA TGTGGACACT ACTCTAAANT AAAAWGTCNA TTTACCTTAA  
--> 1205 AAAAAAATA 600  
1206  
1207 ATCCCCNNTA ANANAAAANT GGGGGGGCCN TTTTGGGGN TCCCGGTTTT  
--> 1208 NGGACGGGGN 660  
1209  
1210 GCTTTNGGGG GGCTTGNNT TTTTNGGN GCCC  
1211  
1212

695

---

1213 (2) INFORMATION FOR SEQ ID NO: 25:  
1214  
1215 (i) SEQUENCE CHARACTERISTICS:  
--> 1216 (A) LENGTH: 495 base pairs  
1217 (B) TYPE: nucleic acid  
1218 (C) STRANDEDNESS: double  
1219 (D) TOPOLOGY: linear  
1220  
1221 (ii) MOLECULE TYPE: cDNA to mRNA  
1222  
1223 (iii) HYPOTHETICAL: NO  
1224  
1225 (iii) ANTI-SENSE: NO  
1226  
1227 (vi) ORIGINAL SOURCE:  
1228 (A) ORGANISM: Arabidopsis thaliana  
1229 (B) STRAIN: ecotype Columbia  
1230  
1231 (ix) FEATURE:  
1232 (A) NAME/KEY: CDS  
1233 (B) LOCATION: 2..495  
1234  
1235 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
1236

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:38

INPUT SET: S33211.raw

--> 1237 TCTGTTTTNA GGCAGAGCAG AGGAAGTTGT TGCTTTGCTT GGTAAGGAGT  
1238 TTCCTGAATT 60  
1239  
1240 NAGTTTAAAG AAGGAGAACT GTTCGGAGAT GACTTGGTTT CAGTCAGCTT  
--> 1241 TATGGTGGGA 120  
1242  
1243 TAATCGTGTT AACCTACTC ANATTGATCC WAAAGTGTTT CTCGATCGGA  
--> 1244 ATCTTGATAG 180  
1245  
1246 AGCGAATTTC GGAAAGAGGA AATCGGATTA CGTTGCGAGT AAGATTCCTA  
--> 1247 GAGATGGGAT 240  
1248  
1249 TAAGYCTTTT TCCAAGARGA TGMCTGACCT GGGGAAAAYC GGGCTTGTTT  
--> 1250 TTAAWCCGTA 300  
1251  
1252 TGGTGGGAAA ATGGCGGAGG TTACGGTTAA CGCGACGCCG TTTCCNCACC  
--> 1253 GAAGCAAGCT 360  
1254  
1255 TTTTAAGATT CAGTACTCGG TGACTTNGCA AGAAAACCTCT NTCGAGATAG  
--> 1256 AGAAAGGGTT 420  
1257  
1258 TCTTGAATCA GGCTAACGTC CTTATAGGTT CATGACCGGG TTTTNNAGCA  
--> 1259 AGANCCCTGG 480  
1260  
1261 AATNCTTACT TNAAT  
1262  
1263

*same*

495

---

1264 (2) INFORMATION FOR SEQ ID NO: 26:  
1265  
1266 (i) SEQUENCE CHARACTERISTICS:  
--> 1267 (A) LENGTH: 204 base pairs  
1268 (B) TYPE: nucleic acid  
1269 (C) STRANDEDNESS: double  
1270 (D) TOPOLOGY: linear  
1271  
1272 (ii) MOLECULE TYPE: cDNA to mRNA  
1273  
1274 (iii) HYPOTHETICAL: NO  
1275  
1276 (iii) ANTI-SENSE: NO  
1277  
1278 (vi) ORIGINAL SOURCE:  
1279 (A) ORGANISM: Arabidopsis thaliana  
1280 (B) STRAIN: ecotype Columbia  
1281  
1282 (ix) FEATURE:  
1283 (A) NAME/KEY: CDS  
1284 (B) LOCATION: 1..204  
1285  
1286 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
1287  
1288 AAATTAAAAC AAATCAATGT TGATATTGAA TCCAATAGTG CTTGGTTTCA

*next page*

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:38

INPUT SET: S33211.raw

--> 1289 ACCTGGTGCT 60  
1290  
1291 ACGCTTGGTG AGCTTTACTA CAGAATTNCA GAGAAGAGCA AAATCCATGG  
--> 1292 ATTTCCNGCG 120  
1293  
1294 GGTTTNTNCA CAAGCNTAGG CATAGGTGGG TATATNANAG GCGGTGGATA  
--> 1295 CCGTACCTTG 180  
1296  
1297 ATGAGGAAGT ATGGTCTTNC GGA  
1298  
1299

204

---

1300 (2) INFORMATION FOR SEQ ID NO: 27:  
1301  
1302 (i) SEQUENCE CHARACTERISTICS:  
--> 1303 (A) LENGTH: 491 base pairs  
1304 (B) TYPE: nucleic acid  
1305 (C) STRANDEDNESS: double  
1306 (D) TOPOLOGY: linear  
1307  
1308 (ii) MOLECULE TYPE: cDNA to mRNA  
1309  
1310 (iii) HYPOTHETICAL: NO  
1311  
1312 (iii) ANTI-SENSE: NO  
1313  
1314 (vi) ORIGINAL SOURCE:  
1315 (A) ORGANISM: Arabidopsis thaliana  
1316 (B) STRAIN: ecotype Columbia  
1317  
1318 (ix) FEATURE:  
1319 (A) NAME/KEY: CDS  
1320 (B) LOCATION: 2..491  
1321  
1322 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
1323  
1324 GAGATTTCTC GAGCAAGATA CTCCACTGAT GATCTTTGAG CCATTGGGTG  
--> 1325 GGAAAATCAG 60  
1326  
1327 CAAGATTTCA GAAACAGAAT CTCCATATCC ACACAGAAGA GGTAATCTGT  
--> 1328 ATAATATACA 120  
1329  
1330 GTACATGGTG AAATGGAAAG TGAATGANGT CGAGGAGATG AACAAACATG  
--> 1331 TCAGGTGGAT 180  
1332  
1333 GAGATCGTTA CACGATTACA TGA CTCCGTA TGTTTCTAAA TCGCCGAGAG  
--> 1334 GAGCTTATTT 240  
1335  
1336 GANTTACAGA GATCTTGATT TGGGCTCGAC CAAAGGGATT AACACGGGTT  
--> 1337 TCGGAGATGC 300  
1338  
1339 AAGGAAATGG NNGGGTGAGN CTTTTTTTCAA AGGTAATTTT CAAGGGGTTA  
--> 1340 GGTTTTGGTT 360



RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:39

INPUT SET: S33211.raw

1341  
1342 AAAGGGGAGG TTTNNCCCAN CAAATTTTTT TTCAGGANCC GGCCANGNTT  
--> 1343 TTCCCCCCCC 420  
1344  
1345 TTTTTTTTNGG NCCCCAATCN AAANCCCCGT TTTAAAAGGG GGGCCATTTC  
--> 1346 NTTTTTTTCA 480  
1347  
1348 NNTTAAAGG G  
1349  
1350

491

---

1351 (2) INFORMATION FOR SEQ ID NO: 28:  
1352  
1353 (i) SEQUENCE CHARACTERISTICS:  
--> 1354 (A) LENGTH: 407 base pairs  
1355 (B) TYPE: nucleic acid  
1356 (C) STRANDEDNESS: double  
1357 (D) TOPOLOGY: linear  
1358  
1359 (ii) MOLECULE TYPE: cDNA to mRNA  
1360  
1361 (iii) HYPOTHETICAL: NO  
1362  
1363 (iii) ANTI-SENSE: NO  
1364  
1365 (vi) ORIGINAL SOURCE:  
1366 (A) ORGANISM: Arabidopsis thaliana  
1367 (B) STRAIN: ecotype Columbia  
1368  
1369 (ix) FEATURE:  
1370 (A) NAME/KEY: CDS  
1371 (B) LOCATION: 3..407  
1372  
1373 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
1374  
1375 ATTTGTTTCGT GAGGTAACT TTGACTTTAG TCAACGGTAC GAAGCCTGGT  
--> 1376 GAGAATACGG 60  
1377  
1378 TTTTAGCGAC TTTTATTGGG ATGTATTTAG GCCGGTCGGA TAAGCTGTTG  
--> 1379 ACCGTNATGA 120  
1380  
1381 ACCGGGATTT CCCGGAGTTG AAGCTGAAGA AAACCGATTN TACCGAGATG  
--> 1382 AGATGGATCG 180  
1383  
1384 ATTCGGTTCT GTTTTGGGAC GATTATCCGG TTGGTACACC GACTTCTGTG  
--> 1385 CTACTAAATC 240  
1386  
1387 CGCTAGTCGC AAAAAAGTTG TTCATGAAAC GAAAATCGGA CTACGTGAAG  
--> 1388 CGTCTNATTT 300  
1389  
1390 TCGAGAACCC GATCTCNNGT TTGATACTCA AGAAATTTGT AGAGGTTNNG  
--> 1391 AAAGTTAAAA 360  
1392

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:39

INPUT SET: S33211.raw

407

1393 TNAATTTGGA TCCGCATTNN GGNANNNATG GTGAAACCCC NNGTTNT  
1394  
1395

---

1396 (2) INFORMATION FOR SEQ ID NO: 29:  
1397  
1398 (i) SEQUENCE CHARACTERISTICS:  
--> 1399 (A) LENGTH: 360 base pairs  
1400 (B) TYPE: nucleic acid  
1401 (C) STRANDEDNESS: double  
1402 (D) TOPOLOGY: linear  
1403  
1404 (ii) MOLECULE TYPE: cDNA to mRNA  
1405  
1406 (iii) HYPOTHETICAL: NO  
1407  
1408 (iii) ANTI-SENSE: NO  
1409  
1410 (vi) ORIGINAL SOURCE:  
1411 (A) ORGANISM: Arabidopsis thaliana  
1412 (B) STRAIN: ecotype Columbia  
1413  
1414 (ix) FEATURE:  
1415 (A) NAME/KEY: CDS  
1416 (B) LOCATION: 3..360  
1417  
1418 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
1419  
1420 ACGGCGTCGT ATTGGCCTAC AAAATAAACC TTGTTGAAGT CCCAGAAAAC  
--> 1421 GTCACCGTTT 60  
1422  
1423 TCAGAATCTC CCGGACGTTA GAACAAAATG CGACGGATAT CATTACCGG  
--> 1424 TGGCAACAAG 120  
1425  
1426 TTGCACCGAA GCTTCCCGAC GAGCTTTTCA TAAGANCAGT CATTGACGTA  
--> 1427 NAAACGGCAC 180  
1428  
1429 TGTTTCATNN CTCAAAAGAC CGTCAGACAA CATTATAGC AATGTTTCTA  
--> 1430 GGAGACACGN 240  
1431  
1432 CAACTCTACT GTCGATATTA AACCGGAGAT TCCCAGAATT GGGTTTGGTC  
--> 1433 CGGTCTGACT 300  
1434  
1435 GTACCGNAAC AAGCNNTTGG ATCCAATCTG TGCTATTTTT GGGACAAATA  
--> 1436 TCCCAGGTTG 360  
1437  
1438

---

1439 (2) INFORMATION FOR SEQ ID NO: 30:  
1440  
1441 (i) SEQUENCE CHARACTERISTICS:  
--> 1442 (A) LENGTH: 427 base pairs  
1443 (B) TYPE: nucleic acid

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:39

INPUT SET: S33211.raw

1444 (C) STRANDEDNESS: double  
1445 (D) TOPOLOGY: linear  
1446  
1447 (ii) MOLECULE TYPE: cDNA to mRNA  
1448  
1449 (iii) HYPOTHETICAL: NO  
1450  
1451 (iii) ANTI-SENSE: NO  
1452  
1453 (vi) ORIGINAL SOURCE:  
1454 (A) ORGANISM: Arabidopsis thaliana  
1455 (B) STRAIN: ecotype Columbia  
1456  
1457 (ix) FEATURE:  
1458 (A) NAME/KEY: CDS  
1459 (B) LOCATION: 3..427  
1460  
1461 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:  
1462  
1463 TCTTCACTGT CACCAAAACG TTAGAACAAG ACGCAAGATT GAAGACTATT  
--> 1464 TCTAAGTGGC 60  
1465  
1466 AACAAATTTC ATCCAAGATT ATTGAAGAGA TACACATCCG AGTGGTACTC  
--> 1467 AGAGCAGCTG 120  
1468  
1469 GAAATGATGG AAACAAGACT GTGACAATGA CCTACCTAGG TCAGTTTCTT  
--> 1470 GGCGAGAAAG 180  
1471  
1472 GCACCTTGCT GAAGGTTATG GAGAAGGCTT TTCCAGAACT AGGGTTAACT  
--> 1473 CAAAAGGATT 240  
1474  
1475 GTACTGAAAT GAGCTGGATT GAAGCCGCCC TTTTCCATGG TGGRITTTCCA  
--> 1476 ACAGGKTCTC 300  
1477  
1478 CTATTGAAAT TTTGCTTMAG CTCAAGTCGC CTYTAGGAAA AGRTTWCTTC  
--> 1479 AAAGCAACGK 360  
1480  
1481 CGGATTTTCGT TAAAGAACCT WTTCTGTGA TAGGGCTCAA AGGAATATTC  
--> 1482 AAAAGATTGA 420  
1483  
1484 TTGAAGG  
1485  
1486

427

---

1487 (2) INFORMATION FOR SEQ ID NO: 31:  
1488  
1489 (i) SEQUENCE CHARACTERISTICS:  
--> 1490 (A) LENGTH: 437 base pairs  
1491 (B) TYPE: nucleic acid  
1492 (C) STRANDEDNESS: double  
1493 (D) TOPOLOGY: linear  
1494  
1495 (ii) MOLECULE TYPE: cDNA to mRNA

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:40

INPUT SET: S33211.raw

1496  
1497 (iii) HYPOTHETICAL: NO  
1498  
1499 (iii) ANTI-SENSE: NO  
1500 (vi) ORIGINAL SOURCE:  
1501 (A) ORGANISM: Arabidopsis thaliana  
1502 (B) STRAIN: ecotype Columbia  
1503  
1504 (ix) FEATURE:  
1505 (A) NAME/KEY: CDS  
1506 (B) LOCATION: 1..437  
1507  
1508 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
1509  
1510 GTTGACTAT CATNGAAGAT TAAGTTAGTC GATGTTCCGT CCACGGTCAC  
--> 1511 CGNGTTTAAA 60  
1512  
1513 GTCCAGAAAC ATNAGGAGAA AGAGGCCGTT AGGNTCATCA ACAAGTGGCA  
--> 1514 GTATGTTGCG 120  
1515  
1516 GATAAGGTCC CTGAAGATCT TTTCATCAGC GCAACGTTGG NGAGATCAAA  
--> 1517 CGGAAACTCT 180  
1518  
1519 GTGCAGGCTT TGTTTACTGG ACTCTATCTT GGNCCGGTGA ATAATNTCTT  
--> 1520 GGCCTTGATG 240  
1521  
1522 GAAGAAAAGT TTCCAGANTT AGGTCTTGAT ATCCAAGNCT GCACAGAGAT  
--> 1523 GAGTTGGGCT 300  
1524  
1525 GAATCTGCAC TCTGGTNTNC TGNTTTCNCT AAAGGAGAGN CTCCTTGGGT  
--> 1526 GTTCCNCGCG 360  
1527  
1528 GATCGGNAGC GGNCAATTTN TGGNCTTTCA AGGGGAAAGN CGGCTTTTTN  
--> 1529 CAAGAACCCG 420  
1530  
1531 NTACCCGGGG TTCAATT  
1532  
1533

437

---

1534 (2) INFORMATION FOR SEQ ID NO: 32:  
1535  
1536 (i) SEQUENCE CHARACTERISTICS:  
--> 1537 (A) LENGTH: 441 base pairs  
1538 (B) TYPE: nucleic acid  
1539 (C) STRANDEDNESS: double  
1540 (D) TOPOLOGY: linear  
1541  
1542 (ii) MOLECULE TYPE: cDNA to mRNA  
1543  
1544 (iii) HYPOTHETICAL: NO  
1545  
1546 (iii) ANTI-SENSE: NO  
1547

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:40

INPUT SET: S33211.raw

1548 (vi) ORIGINAL SOURCE:  
1549 (A) ORGANISM: Arabidopsis thaliana  
1550  
1551 (ix) FEATURE:  
1552 (A) NAME/KEY: CDS  
1553 (B) LOCATION: 1..441  
1554  
1555 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
1556  
1557 GCGGACCCTA TAGATCANNA TGTGCTACTG ANAGAAGAGG AAGCCAAGAA  
--> 1558 CAAGCCGGAG 60  
1559  
1560 ACAGATAAAT ATCTGAAATG GGNCGATANC GTTTACGAAT TTATGACNCC  
--> 1561 ATATGTTTTG 120  
1562  
1563 AAATCTCCAA GAGGAGCTTA TGTCAATTTC AAGGATATGG ATTTGGGTAT  
--> 1564 GTATCTTGGA 180  
1565  
1566 AAGAAGAAGA CAAAGTACGA GGAAGGAAAG AGTTGGGGAG TGAAGTATTT  
--> 1567 CAAGAACAAT 240  
1568  
1569 TTCGAGAGAT TGGTGAGAGT GAAGACTAGG GTTGATCCAA CAGATTTCTT  
--> 1570 CTGCGATGAA 300  
1571  
1572 CAGAGCATTC CTCTGGTGAA CAAAGTTACC TGAAGATATC ATTTGAAGTT  
--> 1573 TTTTATTAGT 360  
1574  
1575 CCCTTTTCTC TGTGAAATCA TCTGTGCGTG TTGAATATTA TGCGTCAAGT  
--> 1576 GTGTAACCTA 420  
1577  
1578 TGTGTGTGAT TGTGAATTGT G  
1579  
1580

441

---

1581 (2) INFORMATION FOR SEQ ID NO: 33:  
1582  
1583 (i) SEQUENCE CHARACTERISTICS:  
--> 1584 (A) LENGTH: 502 base pairs  
1585 (B) TYPE: nucleic acid  
1586 (C) STRANDEDNESS: double  
1587 (D) TOPOLOGY: linear  
1588  
1589 (ii) MOLECULE TYPE: cDNA to mRNA  
1590  
1591 (iii) HYPOTHETICAL: NO  
1592  
1593 (iii) ANTI-SENSE: NO  
1594  
1595 (vi) ORIGINAL SOURCE:  
1596 (A) ORGANISM: Arabidopsis thaliana  
1597 (B) STRAIN: ecotype Columbia  
1598  
1599 (ix) FEATURE:

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:41

INPUT SET: S33211.raw

1600 (A) NAME/KEY: CDS  
1601 (B) LOCATION: 2..502  
1602  
1603 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
1604  
1605 CTGGCTTAAC ACAACGTCGT TTTGGGCCAA TTACCCGGCG GGTACACCCA  
--> 1606 AGAGCATCCT 60  
1607  
1608 TCTAGATAGG CCTCCGACGA ATTCAGTGTC ATTTAAGAGT AAATCGGATT  
--> 1609 TTGTCAAAAA 120  
1610  
1611 ACCAATACCC AAAAAAGGTT TAGAGAAGCT TTGGAAGACA ATGTTTAAAT  
--> 1612 TCAACAGTAG 180  
1613  
1614 CGTCTCGTTG CAATTCAACC CTTACGGTGG AGTGATGGAC CGGATTCCGG  
--> 1615 CAACGGCCAC 240  
1616  
1617 CGCTTTTCCT CATCGGAAAG GAAACTTGTT CAAGGTTCAA TACNCTACGA  
--> 1618 TGTGGTTTGA 300  
1619  
1620 CGCAAACGCC ACACAGAGTA GCCNGGCTAT GATGAATGAG CTTTTTGAGG  
--> 1621 TGGCGGGACC 360  
1622  
1623 GTACGTGNGT CAAGTAAACC CGAGANANGG CTTCTTTTAA NTTTCAGAGNC  
--> 1624 CATCGNTNTT 420  
1625  
1626 NGGAGCAANN CCAAGTGGGG GGGNCCAACC GGGGGNTNAA ANCNNAGNTC  
--> 1627 TTNGGGGGCC 480  
1628  
1629 CAGAATTTCC TTNGGGGAAT TT  
1630  
1631

502

---

1632 (2) INFORMATION FOR SEQ ID NO: 34:  
1633  
1634 (i) SEQUENCE CHARACTERISTICS:  
--> 1635 (A) LENGTH: 400 base pairs  
1636 (B) TYPE: nucleic acid  
1637 (C) STRANDEDNESS: double  
1638 (D) TOPOLOGY: linear  
1639  
1640 (ii) MOLECULE TYPE: cDNA to mRNA  
1641  
1642 (iii) HYPOTHETICAL: NO  
1643  
1644 (iii) ANTI-SENSE: NO  
1645  
1646 (vi) ORIGINAL SOURCE:  
1647 (A) ORGANISM: Arabidopsis thaliana  
1648 (B) STRAIN: ecotype Columbia  
1649  
1650 (ix) FEATURE:  
1651 (A) NAME/KEY: CDS

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:41

INPUT SET: S33211.raw

1652 (B) LOCATION: 2..400  
1653  
1654 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:  
1655  
1656 NGGGAATTGC NCGAGGNAAG TTGTACCCAA TTCCTGGACC ACCATTGGTT  
--> 1657 TCCCAAGAAN 60  
1658  
1659 CCCGAGACAA CCGTTTTTCA ATNACCGTGA TGTTGATTG GGTATTAATT  
--> 1660 CTCATAATGG 120  
1661  
1662 TAAAATCAGT AGTTATGTGG AAGGTAAACG TTACGGGAAG AAGTATTTTCG  
--> 1663 CAGGTAATTT 180  
1664  
1665 CGAGAGATTG GTGAAGATTA AGACGAGAGT TGATAGTGGT AATTTCTTTA  
--> 1666 GGAACGAACA 240  
1667  
1668 GAGTATTCCT GTGTTACCAT AAGTGTATTT ATTTGATTAT TGGTTAGTGA  
--> 1669 AATTTGTTGT 300  
1670  
1671 TGTATAATGA TTATATGTCG TATTTTTATT TATTATTAGT AATTTATAAA  
--> 1672 GTTTGATATT 360  
1673  
1674 AAATACAAAT AGTATAATAA GATAGTTTCT TTTAGTAAAA  
1675  
1676

400

---

1677 (2) INFORMATION FOR SEQ ID NO: 35:  
1678  
1679 (i) SEQUENCE CHARACTERISTICS:  
--> 1680 (A) LENGTH: 383 base pairs  
1681 (B) TYPE: nucleic acid  
1682 (C) STRANDEDNESS: double  
1683 (D) TOPOLOGY: linear  
1684  
1685 (ii) MOLECULE TYPE: cDNA to mRNA  
1686  
1687 (iii) HYPOTHETICAL: NO  
1688  
1689 (iii) ANTI-SENSE: NO  
1690  
1691 (vi) ORIGINAL SOURCE:  
1692 (A) ORGANISM: Arabidopsis thaliana  
1693 (B) STRAIN: ecotype Columbia  
1694  
1695 (ix) FEATURE:  
1696 (A) NAME/KEY: CDS  
1697 (B) LOCATION: 2..383  
1698  
1699 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
1700  
1701 CAACTCTAAT GGGAACACCT ACTTCGATCG AATGTCGATG GGGAAGAGC  
--> 1702 TTTTCTGGGC 60  
1703

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:41

INPUT SET: S33211.raw

--> 1704 GGTTTCGAGGA GGTGGAGCCG CGAGTTTCGG CATCGTGATG GGATACAAAA  
1705 TCCGGTTGGT 120  
1706  
1707 TCCGGTTCCG GAGAAAGTTA CGGTTTTTAG CGTCGGAAAA ACCGTCGGAG  
--> 1708 AAGGAGCCGT 180  
1709  
1710 TGATCTTATA ATGAAGTGGC AGAACTTCTC TCATAGTACG GNTCGGAATT  
--> 1711 TTTTTGTGAA 240  
1712  
1713 GCTGANTTTT GANTTTAGTC AACGGTGCAA AGCCGGGTGA AAAAAAGGTT  
--> 1714 TTAGNGNCTT 300  
1715  
1716 TCANTTTGGN TGNAANCTTG GGGGTTTTAT NAGAACGGTT AACCGGGATT  
--> 1717 NANCCCGNGT 360  
1718  
1719 TTTCCCGGGG TTAAACCTT NGG  
1720  
1721

383

---

1722 (2) INFORMATION FOR SEQ ID NO: 36:  
1723  
1724 (i) SEQUENCE CHARACTERISTICS:  
--> 1725 (A) LENGTH: 354 base pairs  
1726 (B) TYPE: nucleic acid  
1727 (C) STRANDEDNESS: double  
1728 (D) TOPOLOGY: linear  
1729  
1730 (ii) MOLECULE TYPE: cDNA to mRNA  
1731  
1732 (iii) HYPOTHETICAL: NO  
1733  
1734 (iii) ANTI-SENSE: NO  
1735  
1736 (vi) ORIGINAL SOURCE:  
1737 (A) ORGANISM: Arabidopsis thaliana  
1738 (B) STRAIN: ecotype Columbia  
1739  
1740 (ix) FEATURE:  
1741 (A) NAME/KEY: CDS  
1742 (B) LOCATION: 1..354  
1743  
1744 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:  
1745  
1746 ATCAATGTTT TTAATAAACG TACACGAGCA TCGTTGGCTT TCAAGGCTAA  
--> 1747 ATCTGATTTT 60  
1748  
1749 NTTCAAGAAC CGATNCCTAA AACCGCGATT TCGAAGCTTT GGAGACGGTT  
--> 1750 GCAAGAACCG 120  
1751  
1752 GAAGCAGAGC ATGCTCAGCT AATTTCACN CCATTTGGTG GTAAAATGAG  
--> 1753 TNAGATTGCA 180  
1754  
1755 GATTACGAAA CACCATTTCG GCATAGGAAG GGGAATATAT ATNAGATTCA



RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:42

INPUT SET: S33211.raw

--> 1756 GTACTTGAAT 240  
1757  
1758 TACTGGAGAG GAGACGTGAA AGAGAAGTAT ATTGAGATNG GTGGAGGAGA  
--> 1759 GTTTACGGTT 300  
1760  
1761 GNTATNAGTA AGTTTTTTTGG CGAAGTNTNC CNAGAGGNGN CTTNNTNTAA ACCT  
1762 354  
1763  
1764

---

1765 (2) INFORMATION FOR SEQ ID NO: 37:  
1766  
1767 (i) SEQUENCE CHARACTERISTICS:  
--> 1768 (A) LENGTH: 403 base pairs  
1769 (B) TYPE: nucleic acid  
1770 (C) STRANDEDNESS: double  
1771 (D) TOPOLOGY: linear  
1772  
1773 (ii) MOLECULE TYPE: cDNA to mRNA  
1774  
1775 (iii) HYPOTHETICAL: NO  
1776  
1777 (iii) ANTI-SENSE: NO  
1778  
1779 (vi) ORIGINAL SOURCE:  
1780 (A) ORGANISM: Arabidopsis thaliana  
1781 (B) STRAIN: ecotype Columbia  
1782  
1783 (ix) FEATURE:  
1784 (A) NAME/KEY: CDS  
1785 (B) LOCATION: 2..403  
1786  
1787 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:  
1788  
1789 TTTTTTAGTA CACTAATAAT CAAATGGAAT GAGAAATGAA GCCACAAAAG  
--> 1790 TATCTGCAAT 60  
1791  
1792 CAAAATATCC TGCTATCTCC ATCTCAAGCT CTCAATAGTA TCCTCTCCGA  
--> 1793 AAGTGAAATC 120  
1794  
1795 AACATTTCAA ACTCTATTTT TGGTGGAAT CGATAGACTG ATTCCTCTGA  
--> 1796 TGAACCAGAA 180  
1797  
1798 GTTTCGGGAA CTCGGCTTAC GATCTCAAGA CTGTTCGGAA ATGAGCTGGA  
--> 1799 TCGAATCGAT 240  
1800  
1801 AATGTTCTTC AACTGGAGAT CAGGACAGCC GTTAGAGATT TTGCTCAACA  
--> 1802 GAGACCTAAG 300  
1803  
1804 GATTCGAGGA TCAGTATTTT AAAGCAAAGT CAGGATTATG GTTCAAAAAC  
--> 1805 CCGTTCCTGA 360  
1806  
1807 AAACGTTTTT CGAAGAGGTA TCCAAGGGGT TTCTCGAGCA AGT

403

INPUT SET: S33211.raw

1808  
1809

---

1810 (2) INFORMATION FOR SEQ ID NO: 38:  
1811  
1812 (i) SEQUENCE CHARACTERISTICS:  
--> 1813 (A) LENGTH: 260 base pairs  
1814 (B) TYPE: nucleic acid  
1815 (C) STRANDEDNESS: double  
1816 (D) TOPOLOGY: linear  
1817  
1818 (ii) MOLECULE TYPE: cDNA to mRNA  
1819  
1820 (iii) HYPOTHETICAL: NO  
1821  
1822 (iii) ANTI-SENSE: NO  
1823  
1824 (vi) ORIGINAL SOURCE:  
1825 (A) ORGANISM: Arabidopsis thaliana  
1826 (B) STRAIN: ecotype Columbia  
1827 (ix) FEATURE:  
1828 (A) NAME/KEY: CDS  
1829 (B) LOCATION: 1..260  
1830  
1831 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:  
1832  
1833 GAGATGAGTT GGATTAANTC TGTACTCTGG TTTGCTGATT TCCCTAAAGG  
--> 1834 AGAATCTCTT 60  
1835  
1836 NGTGTCTCTCA CGAATCGTAA GCGTACATCT CTATCTTTNA AAGGCAAAGA  
--> 1837 TGATTTTATC 120  
1838  
1839 CAAGAACCGA TACCCGAGGC TGCAATTNAA GAGATATGGA GGCGATTAGA  
--> 1840 AGCCCCCNAG 180  
1841  
1842 GCTCGGCTTG GAAAGATCAT ATTA ACTCCA TTTGGTGGGA AAATNAGTGA  
--> 1843 AATGGCAGAG 240  
1844  
1845 TACGTANCAC CATTCCCACA  
1846  
1847

260

---

1848 (2) INFORMATION FOR SEQ ID NO: 39:  
1849  
1850 (i) SEQUENCE CHARACTERISTICS:  
--> 1851 (A) LENGTH: 605 base pairs  
1852 (B) TYPE: nucleic acid  
1853 (C) STRANDEDNESS: double  
1854 (D) TOPOLOGY: linear  
1855  
1856 (ii) MOLECULE TYPE: cDNA to mRNA  
1857  
1858 (iii) HYPOTHETICAL: NO

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:42

INPUT SET: S33211.raw

1859  
1860 (iii) ANTI-SENSE: NO  
1861  
1862 (vi) ORIGINAL SOURCE:  
1863 (A) ORGANISM: Arabidopsis thaliana  
1864 (B) STRAIN: ecotype Columbia  
1865  
1866 (ix) FEATURE:  
1867 (A) NAME/KEY: CDS  
1868 (B) LOCATION: 2..605  
1869  
1870 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
1871  
1872 CTCTTGCATA TTCGCTGCAA GGATGGGAAA TTCAAAACCA CTCCCTACAA  
--> 1873 TTTTTTGTAT 60  
1874  
1875 TATAGTTTCA GTCTTGTATT TTTAATTCTA TTGCATAACA CCAACTTCTT  
--> 1876 CATCAGCCTC 120  
1877  
1878 CATCCAAGAT CAATTCATAA ACTGTGTCAA AAGAAACACA CATGTTTCTT  
--> 1879 TTCCACTCGA 180  
1880  
1881 GAAAACGTTA TTCACCCCTG CGAAAAACGT CTCTTTGTTC AACCAAGTCC  
--> 1882 TTGANTCGAC 240  
1883  
1884 GGCTCAAAAT CTCCAGTTCT TGGCAAAATC CATGCCTAAA CCGGGRTTCA  
--> 1885 TATTCAGACC 300  
1886  
1887 GATTCAACCAG TCTCAAGTCC AAGSTTCCAT CATTTGTTCA AMGRAACTCG  
--> 1888 GGNTTCATTT 360  
1889  
1890 TNGTGTTTGA NGTGGCGGTC ACGATTTTTCG AGGCCTTTGT NTTTATGTTT  
--> 1891 CACGGTTTGA 420  
1892  
1893 AAAAACCGTT TATATTACTC GGCCTGTCAA ANTTGNANNC AAAATCANAT  
--> 1894 GTTGGATATT 480  
1895  
1896 GNATTCCAAA TAGGTNCTTG GGGTNAACCT GGTGGCTANC GTTTGGTGAG  
--> 1897 CTTTTACTTT 540  
1898  
1899 CAAGAATTTG CANGNGGANG TGCAAAGATT CCATGGGATT TCCCGGGGGG  
--> 1900 TTTNTTGCAC 600  
1901  
1902 AATGT  
1903  
1904

*same*

605

---

1905 (2) INFORMATION FOR SEQ ID NO: 40:  
1906  
1907 (i) SEQUENCE CHARACTERISTICS:  
--> 1908 (A) LENGTH: 464 base pairs  
1909 (B) TYPE: nucleic acid  
1910 (C) STRANDEDNESS: double

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:43

INPUT SET: S33211.raw

1911 (D) TOPOLOGY: linear  
1912  
1913 (ii) MOLECULE TYPE: cDNA to mRNA  
1914  
1915 (iii) HYPOTHETICAL: NO  
1916  
1917 (iii) ANTI-SENSE: NO  
1918  
1919 (vi) ORIGINAL SOURCE:  
1920 (A) ORGANISM: Arabidopsis thaliana  
1921 (B) STRAIN: ecotype Columbia  
1922  
1923 (ix) FEATURE:  
1924 (A) NAME/KEY: CDS  
1925 (B) LOCATION: 2..464  
1926  
1927 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:  
1928  
1929 AACACAAAAC TCTTCCATTT GGCTTCTCTC TTGCATATTC GTTGCAAGGA  
--> 1930 TGGGAAATTC 60  
1931  
1932 AAAACCACTC CCTACAATTN CTTGTATTAT CGTTTCAGTC TTGTATTTTN  
--> 1933 NATTCTATTG 120  
1934  
1935 CATAACACCA ACTTCTTCAT CAGCCTCCAT CCAAGNTCAA TTCATAAACT  
--> 1936 GTGTCAAAAG 180  
1937  
1938 GAACACACAT GTTTCTTTTC CACTCGAGNA AACGGTATTC ACTCCTGCGG  
--> 1939 AAAACGGCTC 240  
1940  
1941 TTTTATTCAA CGGGTCCNTG AATCGACGGG TCAAAATCTC CAGTTCTTGG  
--> 1942 NAAAATCCAT 300  
1943  
1944 GNCTAAACCG GGGTTCATAT TCAGGCCGGT TCACCAGTCT CAAGTCCAAG  
--> 1945 NTTCCATCAT 360  
1946  
1947 TTGTTCAAAG GAACTCGGGA TTCATTTCCG CGNTAGAAGT GGCGGGCANN  
--> 1948 GGTTTCGGGG 420  
1949  
1950 CCTGTCTNTT GNTTANGGGN AGGAAAACCG GTTNTATTNC TCGG  
1951  
1952

464

---

1953 (2) INFORMATION FOR SEQ ID NO: 41:  
1954  
1955 (i) SEQUENCE CHARACTERISTICS:  
--> 1956 (A) LENGTH: 386 base pairs  
1957 (B) TYPE: nucleic acid  
1958 (C) STRANDEDNESS: double  
1959 (D) TOPOLOGY: linear  
1960  
1961 (ii) MOLECULE TYPE: cDNA to mRNA  
1962

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:43

INPUT SET: S33211.raw

1963 (iii) HYPOTHETICAL: NO  
1964  
1965 (iii) ANTI-SENSE: NO  
1966  
1967 (vi) ORIGINAL SOURCE:  
1968 (A) ORGANISM: Arabidopsis thaliana  
1969 (B) STRAIN: ecotype Columbia  
1970  
1971 (ix) FEATURE:  
1972 (A) NAME/KEY: CDS  
1973 (B) LOCATION: 1..386  
1974  
1975 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:  
1976  
1977 TCGGGAGCCC ANGNTAAATT ANNTGAAAAT GGGGNCGNAT ANCCGTTTAC  
--> 1978 NGAATTTTAT 60  
1979  
1980 GACNCCCAAT ATGTTTCGAA ATCTCAAAGA NNGGGANCTT ATGTCAATTT  
--> 1981 CAAGGATATG 120  
1982  
1983 GATTTGGGTA TGTATCTTGG AAAGNAGAAG ACAAAGTACG AGGAAGGAAA  
--> 1984 GAGTTGGGGA 180  
1985  
1986 GTGAAGTATT TCAAGAACAA TTTCGAGAGA TTGGTGAGAG TGAAGACTAG  
--> 1987 GGTGATCCN 240  
1988  
1989 ACAGATTTTCN TCTGCGATGA ACAGAGCATT CCTCTGGTGN ACAAAGTTAC  
--> 1990 CTGAAGATAT 300  
1991  
1992 CATTTGAAGT TTTTATTAG TCCCTTTTCT CTGTGAAATC ATCTGTGCGT  
--> 1993 GTTGAATANT 360  
1994  
1995 ATGCGTCAAG TGTGTAACCTT ATGTGT  
1996  
1997

386

---

1998 (2) INFORMATION FOR SEQ ID NO: 42:  
1999  
2000 (i) SEQUENCE CHARACTERISTICS:  
--> 2001 (A) LENGTH: 377 base pairs  
2002 (B) TYPE: nucleic acid  
2003 (C) STRANDEDNESS: double  
2004 (D) TOPOLOGY: linear  
2005  
2006 (ii) MOLECULE TYPE: cDNA to mRNA  
2007  
2008 (iii) HYPOTHETICAL: NO  
2009  
2010 (iii) ANTI-SENSE: NO  
2011  
2012 (vi) ORIGINAL SOURCE:  
2013 (A) ORGANISM: Arabidopsis thaliana  
2014 (B) STRAIN: ecotype Columbia

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:44

INPUT SET: S33211.raw

2015  
2016 (ix) FEATURE:  
2017 (A) NAME/KEY: CDS  
2018 (B) LOCATION: 1..377  
2019  
2020 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:  
2021  
2022 TACCATAGGG AGGTGGTGNA AGATTTTGTA TGTAGNCTTA GGGGAAGGCG  
--> 2023 AGTAGTATGG 60  
2024 TGGTGGTGGG GAGCTGTAAA CGTATGGTGG TGGTGGAGAT TTGTATGTGG  
--> 2025 GCTGGTTAAC 120  
2026  
2027 TTCATTGAAG CTAAAATCTG GGGACCTAAG TACTTCAAAG GCAATTTTGA  
--> 2028 CAGATTGGTG 180  
2029  
2030 AAGATTAAAA CCAAGGTTGA TCCAGAGAAC TTCTTCAGGC ACGAGCAGAG  
--> 2031 TATCCACCT 240  
2032  
2033 ATGCCCTACT AGAAGCTAGG TTCATGAAAC CAATAACATT ATCAAAAATA  
--> 2034 AGAATAAATG 300  
2035  
2036 ATAATTGTAT ACAACATGAT TCGTCTTTCT TTATTTTCTA CAATGTGGAC  
--> 2037 ACTACTCTAA 360  
2038  
2039 ATAAATGTC ATTTACC  
2040  
2041

377

---

2042 (2) INFORMATION FOR SEQ ID NO: 43:  
2043  
2044 (i) SEQUENCE CHARACTERISTICS:  
--> 2045 (A) LENGTH: 377 base pairs  
2046 (B) TYPE: nucleic acid  
2047 (C) STRANDEDNESS: double  
2048 (D) TOPOLOGY: linear  
2049  
2050 (ii) MOLECULE TYPE: cDNA to mRNA  
2051  
2052 (iii) HYPOTHETICAL: NO  
2053  
2054 (iii) ANTI-SENSE: NO  
2055  
2056 (vi) ORIGINAL SOURCE:  
2057 (A) ORGANISM: Arabidopsis thaliana  
2058 (B) STRAIN: ecotype Columbia  
2059  
2060 (ix) FEATURE:  
2061 (A) NAME/KEY: CDS  
2062 (B) LOCATION: 1..377  
2063  
2064 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:  
2065  
2066 TACCATAGGG AGGTGGTGNA AGATTTTGTA TGTAGNCTTA GGGGAAGGCG

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:44

INPUT SET: S33211.raw

--> 2067 AGTAGTATGG 60  
2068  
2069 TGGTGGTGGG GAGCTGTAAA CGTATGGTGG TGGTGGAGAT TTGTATGTGG  
--> 2070 GCTGGTTAAC 120  
2071  
2072 TTCATTGAAG CTAAAATCTG GGGACCTAAG TACTTCAAAG GCAATTTTGA  
--> 2073 CAGATTGGTG 180  
2074  
2075 AAGATTAAAA CCAAGGTTGA TCCAGAGAAC TTCTTCAGGC ACGAGCAGAG  
--> 2076 TATCCACCT 240  
2077  
2078 ATGCCCTACT AGAAGCTAGG TTCATGAAAC CAATAACATT ATCAAAAATA  
--> 2079 AGAATAAATG 300  
2080  
2081 ATAATTGTAT ACAACATGAT TCGTCTTTCT TTATTTTCAGA CAATGTGGAC  
--> 2082 ACTACTCTAA 360  
2083  
2084 ATAAATGTC ATTTACC  
2085  
2086

377

---

2087 (2) INFORMATION FOR SEQ ID NO: 44:  
2088  
2089 (i) SEQUENCE CHARACTERISTICS:  
--> 2090 (A) LENGTH: 346 base pairs  
2091 (B) TYPE: nucleic acid  
2092 (C) STRANDEDNESS: double  
2093 (D) TOPOLOGY: linear  
2094  
2095 (ii) MOLECULE TYPE: cDNA to mRNA  
2096  
2097 (iii) HYPOTHETICAL: NO  
2098  
2099 (iii) ANTI-SENSE: NO  
2100  
2101 (vi) ORIGINAL SOURCE:  
2102 (A) ORGANISM: Arabidopsis thaliana  
2103 (B) STRAIN: ecotype Columbia  
2104  
2105 (ix) FEATURE:  
2106 (A) NAME/KEY: CDS  
2107 (B) LOCATION: 2..346  
2108  
2109 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:  
2110  
2111 GAGCTGTGGA TATGGTCACA AATGGCAATC GGTTGGTCCG AAAACTGATC  
--> 2112 CGAATCTTTT 60  
2113  
2114 TATGAGAATN TTGATTCAAC CAGTGACGAG GAAGAAGGTA AAGACTGTGA  
--> 2115 GAGCTTCTNT 120  
2116  
2117 GGTGGCCCTN TTTTNAGGCN AGACAGATGA AGTTTTTGCT TTCCTTAGTA  
--> 2118 AGGAGTTTCC 180

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:44

INPUT SET: S33211.raw

2119  
2120 TGAATTGGGT TTAAAGAAGG AGAATTNTTC GGAGATGACT TGGTTTCANT  
--> 2121 CTGCTTTATG 240  
2122  
2123 GTGGGACAAT CGTCTTAATG CTACTCAGGT TGATCCTAAA GTNTTCTTG  
--> 2124 ATCGGAATCT 300  
2125  
2126 CGATACCTCG AGTTTCGGTA AGAGGAAATC GGATTACGTC GCGACT  
2127  
2128

346

---

2129 (2) INFORMATION FOR SEQ ID NO: 45:  
2130  
2131 (i) SEQUENCE CHARACTERISTICS:  
--> 2132 (A) LENGTH: 261 base pairs  
2133 (B) TYPE: nucleic acid  
2134 (C) STRANDEDNESS: double  
2135 (D) TOPOLOGY: linear  
2136  
2137 (ii) MOLECULE TYPE: cDNA to mRNA  
2138  
2139 (iii) HYPOTHETICAL: NO  
2140  
2141 (iii) ANTI-SENSE: NO  
2142  
2143 (vi) ORIGINAL SOURCE:  
2144 (A) ORGANISM: Arabidopsis thaliana  
2145 (B) STRAIN: ecotype Columbia  
2146  
2147 (ix) FEATURE:  
2148 (A) NAME/KEY: CDS  
2149 (B) LOCATION: 2..261  
2150  
2151 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:  
2152  
2153 ATGGGGTGAG ACTTATTTCA AAGGTAATTT CAAGAGATTA GGTTCGTTA  
--> 2154 AAGGGAAGNT 60  
2155  
2156 TGATCCAACA AATTTCTTCA GGAACGAACA GAGTATTCCT CCTCTGTTTT  
--> 2157 GAGTCCTCAA 120  
2158  
2159 TACAAAACCA GATATAAAAG ATGTCATTTC ATTTTTTCAA TTATAATAGA  
--> 2160 TAATGTAACT 180  
2161  
2162 TTCTGCTACA ATTGTAAAAG TGAGATGTAC CCAATACGGT TTAAGCGGAC  
--> 2163 CGAGAATAGT 240  
2164  
2165 CAATTCAAAG ACCAAATTCT G  
2166  
2167

261

---

2168 (2) INFORMATION FOR SEQ ID NO: 46:  
2169



RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:45

INPUT SET: S33211.raw

2170 (i) SEQUENCE CHARACTERISTICS:  
--> 2171 (A) LENGTH: 478 base pairs  
2172 (B) TYPE: nucleic acid  
2173 (C) STRANDEDNESS: double  
2174 (D) TOPOLOGY: linear  
2175  
2176 (ii) MOLECULE TYPE: cDNA to mRNA  
2177  
2178 (iii) HYPOTHETICAL: NO  
2179  
2180 (iii) ANTI-SENSE: NO  
2181  
2182 (vi) ORIGINAL SOURCE:  
2183 (A) ORGANISM: Arabidopsis thaliana  
2184 (B) STRAIN: ecotype Columbia  
2185  
2186 (ix) FEATURE:  
2187 (A) NAME/KEY: CDS  
2188 (B) LOCATION: 1..478  
2189  
2190 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:  
2191  
2192 GCTCAAAGGA CTAACCATGA AAACCTCCTC AAGTGTCTCT CTCACCGANT  
--> 2193 CAACGAGGAC 60  
2194  
2195 GACTCAAGAN TTATACACAC ATCAAAAGAT CCTTCGTATT TNTCAATCTT  
--> 2196 GATTTCTTCC 120  
2197  
2198 ATACAAAATC CAAGTTTCTC TGTTCCTGAA ACACCTAAAC CGGTTTCAAT  
--> 2199 CATCACTCCG 180  
2200  
2201 GTTCAAGCCA CCGATGTTCA ATCTACGNTT AAATNCGCAC GGNCTTCACG  
--> 2202 GGTATACACA 240  
2203  
2204 ATCAGGGCTA GGAGTGGTNG TCATGACTAC GGAGGTTTAT CTTTACATTG  
--> 2205 GCTTAAAAAN 300  
2206  
2207 CANNCCGTTT GTTNNTCATT GATTTNNAGA AATCTTCCGG GCTTATTTAA  
--> 2208 CATNTAAGAT 360  
2209  
2210 GTTTGATAAN CCGGNCCNG TTTGGGGTTC AAATCCCGGT GGCTTACAAA  
--> 2211 NTTNGGGGGA 420  
2212  
2213 ATTGTNCCTA TGAGGTTTGG AAAATTAANG CAAAATNTTT TGGGCCTTCC  
--> 2214 CGGCCGGT 478  
2215  
2216

---

2217 (2) INFORMATION FOR SEQ ID NO: 47:  
2218  
2219 (i) SEQUENCE CHARACTERISTICS:  
--> 2220 (A) LENGTH: 579 base pairs  
2221 (B) TYPE: nucleic acid

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:45

INPUT SET: S33211.raw

2222 (C) STRANDEDNESS: double  
2223 (D) TOPOLOGY: linear  
2224  
2225 (ii) MOLECULE TYPE: cDNA to mRNA  
2226  
2227 (iii) HYPOTHETICAL: NO  
2228  
2229 (iii) ANTI-SENSE: NO  
2230  
2231 (vi) ORIGINAL SOURCE:  
2232 (A) ORGANISM: Arabidopsis thaliana  
2233 (B) STRAIN: ecotype Columbia  
2234  
2235 (ix) FEATURE:  
2236 (A) NAME/KEY: CDS  
2237 (B) LOCATION: 2..579  
2238  
2239 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:  
2240  
2241 GGCCGTTAGG ATCATCAAGA AATGGCAATA TGCTGCAGAT AAGGTTCTCTG  
--> 2242 ATGATCTTTT 60  
2243  
2244 CATTAGGACA ACATTGGAGA GATCAAACAA GAACGCAGTA CACGCTTTGT  
--> 2245 TCACTGGACT 120  
2246  
2247 ATATATTGGT CCGGTGAACA ATCTATTGGC GTTGATGGAA GAAAAGTTTC  
--> 2248 CGGAACTAGG 180  
2249  
2250 TCTTGAGAAA GAAGGTTGTG AAGAGATGAG TTGGATTGAG TCTGTACTCT  
--> 2251 GGTTCGCTGA 240  
2252  
2253 TTTCCCTAAA GGAGAATCTC TTGGTGTCT CACGAATCGT GAGCGTACAT  
--> 2254 CTCTATCTTT 300  
2255  
2256 CAAAGGCAAA GATGATTTTG TCCAAGAACC GATACCCGAG GCTGCAATTC  
--> 2257 AAGAGATATG 360  
2258  
2259 GAGGCGATTA GAAGCCCCCG AGGCTCGGCT TGGAAAGATC ATATTAATC  
--> 2260 CATTTGGGTG 420  
2261  
2262 NGGNAAAATG AGTGAAATGG CAGAGNCCGA ACCACCAATT CCCACANNCG  
--> 2263 AGGGAGGGGA 480  
2264  
2265 ACCCCTNTGN GGNTCAGAAT GTGGTTCCTG GNNNNNAAGN GGGNGCCAGN  
--> 2266 ACCAANCCGG 540  
2267  
2268 GNCNGTAAAN CNTGNAATGG GCCNAACCCG TNCCGGATT  
2269  
2270

*same*

579


---

2271 (2) INFORMATION FOR SEQ ID NO: 48:  
2272  
2273 (i) SEQUENCE CHARACTERISTICS:

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:45

INPUT SET: S33211.raw

--> 2274 (A) LENGTH: 252 base pairs  
2275 (B) TYPE: nucleic acid  
2276 (C) STRANDEDNESS: double  
2277 (D) TOPOLOGY: linear  
2278  
2279 (ii) MOLECULE TYPE: cDNA to mRNA  
2280  
2281 (iii) HYPOTHETICAL: NO  
2282  
2283 (iii) ANTI-SENSE: NO  
2284  
2285 (vi) ORIGINAL SOURCE:  
2286 (A) ORGANISM: Oryza sativa  
2287 (B) STRAIN: Nipponbare, subsp. japonica  
2288 (D) DEVELOPMENTAL STAGE: etiolated shoot (8 days old)  
2289  
2290 (ix) FEATURE:  
2291 (A) NAME/KEY: CDS  
2292 (B) LOCATION: 3..252  
2293  
2294 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
2295  
2296 TGTCCTGGAA GGTCCGCCTC GTGCAGGTTN CGACGACGGT GACGGTGTTC  
--> 2297 GTCGTCGGGA 60  
2298  
2299 GGAACGTCGA CCAGGGCGCC GCNGACGTCG TCGCCAGATG GCAAGACGTC  
--> 2300 GCGCCGAGCC 120  
2301  
2302 TCCCTCCCGA GCTCACCATA CGGGTGATCG TNCGAGGGCA GCGCGCCACG  
--> 2303 TTCCAGTCGC 180  
2304  
2305 TGTACCTCGG CTCGTGCGCC GACCTGGTGC CGACGATGAG CAGCATGTTC  
--> 2306 CCGGAGCTCG 240  
2307  
2308 GGATGACGAT TG  
2309  
2310



252

---

2575 (2) INFORMATION FOR SEQ ID NO: 57:  
2576  
2577 (i) SEQUENCE CHARACTERISTICS:  
--> 2578 (A) LENGTH: 1981 base pairs  
2579 (B) TYPE: nucleic acid  
2580 (C) STRANDEDNESS: double  
2581 (D) TOPOLOGY: linear  
2582  
2583 (ii) MOLECULE TYPE: cDNA to mRNA  
2584  
2585 (iii) HYPOTHETICAL: NO  
2586  
2587 (iii) ANTI-SENSE: NO  
2588  
2589 (vi) ORIGINAL SOURCE:

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:46

INPUT SET: S33211.raw

2590 (A) ORGANISM: Lactuca sativa  
2591 (B) STRAIN: lollo bionda  
2592  
2593 (ix) FEATURE:  
2594 (A) NAME/KEY: CDS  
2595 (B) LOCATION: 7..1626  
2596  
2597 (ix) FEATURE:  
2598 (A) NAME/KEY: unsure  
2599 (B) LOCATION: replace(372, "g")  
2600  
2601 (ix) FEATURE:  
2602 (A) NAME/KEY: unsure  
2603 (B) LOCATION: replace(379, "g")  
2604  
2605 (ix) FEATURE:  
2606 (A) NAME/KEY: unsure  
2607 (B) LOCATION: replace(786, "t")  
2608  
2609 (ix) FEATURE:  
2610 (A) NAME/KEY: unsure  
2611 (B) LOCATION: replace(1105..1106, "ga")  
2612 (D) OTHER INFORMATION: /note= "also possible "gg" and "aa"  
2613  
2614 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:  
2615  
2616 ACAAAA ATG GCA ATT ACC TAT TCT TTC AAC TTC AAA TCT TAT ATT TTT 48  
2617 Met Ala Ile Thr Tyr Ser Phe Asn Phe Lys Ser Tyr Ile Phe  
2618 1 5 10  
2619  
2620 CCT CTC CTC CTT GTC TTG CTC TCT ACC CAT TCA TCA GCG ACT TCA ACT 96  
2621 Pro Leu Leu Leu Val Leu Leu Ser Thr His Ser Ser Ala Thr Ser Thr  
2622 15 20 25 30  
2623  
2624 TCC ATT ATA GAT CGC TTC ACC CAA TGT CTA AAC AAC CGA GCT GAC CCT 144  
2625 Ser Ile Ile Asp Arg Phe Thr Gln Cys Leu Asn Asn Arg Ala Asp Pro  
2626 35 40 45  
2627  
2628 TCT TTC CCG CTC AGT GGA CAA CTT TAC ACT CCC GAT AAC TCC TCT TTT 192  
2629 Ser Phe Pro Leu Ser Gly Gln Leu Tyr Thr Pro Asp Asn Ser Ser Phe  
2630 50 55 60  
2631  
2632 CCA TCC GTC TTG CAA GCT TAC ATC CGG AAC CTC CGA TTC AAT GAA TCC 240  
2633 Pro Ser Val Leu Gln Ala Tyr Ile Arg Asn Leu Arg Phe Asn Glu Ser  
2634 65 70 75  
2635  
2636 ACG ACT CCC AAA CCC ATC TTA ATC ATC ACC GCC TTA CAC CCT TCA CAC 288  
2637 Thr Thr Pro Lys Pro Ile Leu Ile Ile Thr Ala Leu His Pro Ser His  
2638 80 85 90  
2639  
2640 ATT CAA GCA GCT GTT GTG TGC GCC AAA ACA CAC CGC CTG CTA ATG AAA 336  
2641 Ile Gln Ala Ala Val Cys Ala Lys Thr His Arg Leu Leu Met Lys  
2642 95 100 105 110

since location 372  
can also be "G", use "S"  
(per sequence rules)  
"C" can only  
represent itself

use "R"

use "Y"

use "R" "R"

2643																		
2644	ACC	AGA	AGC	GGA	GGC	CAT	GAT	TAT	GAG	GGG	CTT	TCC	TAT	GTG	ACC	AAT	384	
2645	Thr	Arg	Ser	Gly	Gly	His	Asp	Tyr	Glu	Gly	Leu	Ser	Tyr	Val	Thr	Asn		
2646					115					120					125			
2647																		
2648	TCG	AAC	CAA	CCC	TTT	TTT	GTT	GTT	GAC	ATG	TTC	AAC	TTA	CGC	TCC	ATA	432	
2649	Ser	Asn	Gln	Pro	Phe	Phe	Val	Val	Asp	Met	Phe	Asn	Leu	Arg	Ser	Ile		
2650				130					135					140				
2651																		
2652	AAC	GTG	AGT	ATT	GAA	GAT	GAA	ACT	GCA	TGG	GTC	CAA	GCT	GGT	GCG	ACT	480	
2653	Asn	Val	Ser	Ile	Glu	Asp	Glu	Thr	Ala	Trp	Val	Gln	Ala	Gly	Ala	Thr		
2654			145					150					155					
2655																		
2656	CTT	GGT	GAA	GTC	TAC	TAC	CGA	ATA	GCA	GAG	AAA	AGC	AAC	AGT	CAT	GCT	528	
2657	Leu	Gly	Glu	Val	Tyr	Tyr	Arg	Ile	Ala	Glu	Lys	Ser	Asn	Ser	His	Ala		
2658		160					165					170						
2659																		
2660	TTT	CCG	GCT	GGC	GTT	TGC	CCT	ACT	GTT	GGA	GTT	GGT	GGC	CAT	TTT	AGT	576	
2661	Phe	Pro	Ala	Gly	Val	Cys	Pro	Thr	Val	Gly	Val	Gly	Gly	His	Phe	Ser		
2662	175					180					185					190		
2663																		
2664	GGT	GGT	GGT	TAT	GGT	AAC	TTG	ATG	GGA	AAA	TAC	GGC	CTT	TCT	GTT	GAC	624	
2665	Gly	Gly	Gly	Tyr	Gly	Asn	Leu	Met	Gly	Lys	Tyr	Gly	Leu	Ser	Val	Asp		
2666					195					200					205			
2667																		
2668	AAT	ATT	GTC	GAT	GCT	CAG	TTA	ATC	GAT	GTG	AAT	GGT	AAA	CTT	CTG	AAT	672	
2669	Asn	Ile	Val	Asp	Ala	Gln	Leu	Ile	Asp	Val	Asn	Gly	Lys	Leu	Leu	Asn		
2670				210					215					220				
2671																		
2672	CGG	AAA	TCA	ATG	GGT	GAA	GAT	CTT	TTT	TGG	GCC	ATC	ACA	GGT	GGT	GGT	720	
2673	Arg	Lys	Ser	Met	Gly	Glu	Asp	Leu	Phe	Trp	Ala	Ile	Thr	Gly	Gly	Gly		
2674			225					230					235					
2675																		
2676	GGT	GTC	AGC	TTT	GGT	GTG	GTT	GTA	GCG	TAC	AAG	ATC	AAA	CTG	GTT	CGT	768	
2677	Gly	Val	Ser	Phe	Gly	Val	Val	Val	Ala	Tyr	Lys	Ile	Lys	Leu	Val	Arg		
2678		240					245					250						
2679																		
2680	GTT	CCT	ACC	ACT	GTG	ACC	GTT	TTT	AAC	GTA	CAA	AGA	ACA	TCC	GAG	CAG	816	
2681	Val	Pro	Thr	Thr	Val	Thr	Val	Phe	Asn	Val	Gln	Arg	Thr	Ser	Glu	Gln		

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:47

INPUT SET: S33211.raw

2696 AAC TCT ACC GCT CTT GTT GCC CTC CTG AAC AAG GAT TTC CCT GAA TTA 1008  
2697 Asn Ser Thr Ala Leu Val Ala Leu Leu Asn Lys Asp Phe Pro Glu Leu  
2698 320 325 330  
2699  
2700 GGT GTA GAA ATT TCA GAT TGT ATT GAA ATG AGT TGG ATC GAG TCT GTT 1056  
2701 Gly Val Glu Ile Ser Asp Cys Ile Glu Met Ser Trp Ile Glu Ser Val  
2702 335 340 345 350  
2703  
2704  
2705  
2706 CTT TTC TAC ACA AAC TTC CCC ATT GGT ACT CCG ACC ACT GCT CTT CTA 1104  
2707 Leu Phe Tyr Thr Asn Phe Pro Ile Gly Thr Pro Thr Thr Ala Leu Leu  
2708 355 360 365  
2709  
2710 AGC CGT ACA CCT CAA AGA CTA AAC CCA TTC AAA ATC AAA TCT GAT TAC  
2711 1152  
2712 Ser Arg Thr Pro Gln Arg Leu Asn Pro Phe Lys Ile Lys Ser Asp Tyr  
2713 370 375 380  
2714  
2715 GTA AAA AAC ACT ATT TCC AAA CAG GGA TTC GAA TCC ATA TTT GAA AGG  
2716 1200  
2717 Val Lys Asn Thr Ile Ser Lys Gln Gly Phe Glu Ser Ile Phe Glu Arg  
2718 385 390 395  
2719  
2720 ATG AAA GAA CTC GAA AAC CAA ATG CTA GCT TTC AAC CCT TAT GGT GGA  
2721 1248  
2722 Met Lys Glu Leu Glu Asn Gln Met Leu Ala Phe Asn Pro Tyr Gly Gly  
2723 400 405 410  
2724  
--> 2725 AGA ATG AGC GAA ATT TCC GAA TTT GCA AAG CCT TTT CCC CAT CGA TCA 1296  
2726 Arg Met Ser Glu Ile Ser Glu Phe Ala Lys Pro Phe Pro His Arg Ser  
2727 415 420 425 430  
2728  
2729 GGG AAT ATA GCG AAG ATC CAA TAC GAA GTA AAC TGG GAT GAA CTT GGC  
2730 1344  
2731 Gly Asn Ile Ala Lys Ile Gln Tyr Glu Val Asn Trp Asp Glu Leu Gly  
2732 435 440 445  
2733  
--> 2734 GTT GAA GCA GCC AAT CGG TAC TTG AAC TTC ACA AGG GTG ATG TAT GAT 1392  
2735 Val Glu Ala Ala Asn Arg Tyr Leu Asn Phe Thr Arg Val Met Tyr Asp  
2736 450 455 460  
2737  
2738 TAT ATG ACT CCG TTT GTT TCT AAG AAC CCC AGG GAA GCA TTT CTG AAC 1440  
2739 Tyr Met Thr Pro Phe Val Ser Lys Asn Pro Arg Glu Ala Phe Leu Asn  
2740 465 470 475  
2741  
2742 TAC AGG GAT TTA GAT ATT GGT GTC AAC AGT CAT GGC AAG AAT GCT TAC 1488  
2743 Tyr Arg Asp Leu Asp Ile Gly Val Asn Ser His Gly Lys Asn Ala Tyr  
2744 480 485 490  
2745  
2746 GGT GAA GGA ATG GTT TAT GGG CAC AAG TAT TTC AAA GAG ACG AAT TAT  
2747 1536  
2748 Gly Glu Gly Met Val Tyr Gly His Lys Tyr Phe Lys Glu Thr Asn Tyr

1152  
format  
error

←

←

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:47

INPUT SET: S33211.raw

```

2749      495              500              505              510
2750
--> 2751  AAG AGG CTA ACG ATG GTG AAG ACG AGG GTT GAT CCT AGC AAT TTT TTT 1584
2752      Lys Arg Leu Thr Met Val Lys Thr Arg Val Asp Pro Ser Asn Phe Phe
2753              515              520              525
2754
2755  AGG AAT GAG CAA AGT ATC CCA ACT TTG TCA TCT TCA TGG AAG      1626
2756      Arg Asn Glu Gln Ser Ile Pro Thr Leu Ser Ser Ser Trp Lys
2757              530              535              540
2758
2759  TAAATTCTAA ATTCACCTTGT GAAATTGAAT AAAAGTATGG CTTTTTCAAG
--> 2760  GTCATGGTAT 1686
2761
2762  CCAGATTCAG ATGATATTGA TATAATTTTG ACTTGTATTT ATACAAACAA
--> 2763  AATTATATTA 1746
2764
2765  TATTTTTCTG AATTTAGATT TTCCATTCTT TGGAAAAATA TACGAACATT
--> 2766  GATGTTGATA 1806
2767
2768  TTTTTAAGAA TTATAGATTT TGAACATTGT GAACAATGAA TAAACCGAGG
--> 2769  ACTTCCCTTG 1866
2770
2771  GGTTTTTTTT ATAAGTATGT AATAGCATGT CTTTAATCAA GATAACCGAT
--> 2772  CATTGGATGC 1926
2773
2774  AATTTATTAT TATAAACCTT ATTTAAAAAA AAAAAAAAAA AAAAAAAAAA
--> 2775  AAAAA 1981
2776
2777

```

*same*

---

```

3081  (2) INFORMATION FOR SEQ ID NO: 70:
3082
3083      (i) SEQUENCE CHARACTERISTICS:
--> 3084          (A) LENGTH: 1757 base pairs
3085          (B) TYPE: nucleic acid
3086          (C) STRANDEDNESS: double
3087          (D) TOPOLOGY: linear
3088
3089      (ii) MOLECULE TYPE: DNA (genomic)
3090
3091      (iii) HYPOTHETICAL: NO
3092
3093      (iii) ANTI-SENSE: NO
3094
3095      (vi) ORIGINAL SOURCE:
3096          (A) ORGANISM: Arabidopsis thaliana
3097          (B) STRAIN: Colombia
3098
3099      (ix) FEATURE:
3100          (A) NAME/KEY: CDS
3101          (B) LOCATION: join(1..570, 801..1754)
3102
3103      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

```

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

 DATE: 09/03/1999  
 TIME: 14:51:47

INPUT SET: S33211.raw

```

3104
3105 ACT TCC CGT AGA AAC TCG GAG ACT TTC ACA CAA TGC CTA ACC TCA AAC 48
3106 Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn
3107 1 5 10 15
3108
3109 TCC GAC CCC AAA CAT CCC ATC TCC CCC GCT ATC TTC TTC TCC GGA AAT 96
3110 Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn
3111 20 25 30
3112
3113 GGC TCC TAC TCC TCC GTA TTA CAA GCC AAC ATC CGT AAC CTC CGC TTC 144
3114 Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe
3115 35 40 45
3116
3117 AAC ACC ACC TCA ACT CCG AAA CCC TTC CTC ATA ATC GCC GCA ACA CAT 192
3118 Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His
3119 50 55 60
3120
3121 GAA TCC CAT GTG CAA GCC GCG ATT ACT TGC GGG AAA CGC CAC AAC CTT 240
3122 Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu
3123 65 70 75 80
3124
3125 CAG ATG AAA ATC AGA AGT GGA GGC CAC GAC TAC GAT GGC TTG TCA TAC
3126 288
3127 Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr
3128 85 90 95
3129
--> 3130 GTT ACA TAC TCT GGC AAA CCG TTC TTC GTC CTC GAC ATG TTT AAC CTC 336
3131 Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu
3132 100 105 110
3133
3134 CGT TCG GTG GAT GTC GAT GTG GCA AGT AAG ACC GCG TGG GTC CAA ACC 384
3135 Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr
3136 115 120 125
3137
3138 GGT GCC ATA CTC GGA GAA GTT TAT TAC TAT ATA TGG GAG AAG AGC AAA 432
3139 Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys
3140 130 135 140
3141
3142 ACC CTA GCT TAT CCC GCC GGA ATT TGT CCC ACG GTT GGT GTC GGT GGC 480
3143 Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly
3144 145 150 155 160
3145
3146 CAT ATC AGT GGT GGA GGT TAC GGT AAC ATG ATG AGA AAA TAC GGT CTC 528
3147 His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu
3148 165 170 175
3149
3150 ACC GTA GAT AAT ACC ATC GAT GCA AGA ATG GTC GAC GTT AAT 570
3151 Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn
3152 180 185 190
3153
--> 3154 GGTATAATTG ATATCTCTAT TTTATATACT AATTAAATTT TATAGTGTGG
3155 ATCGGATAGT 630
3156

```

*same**same*



# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:48

INPUT SET: S33211.raw

```

3157 GATTTTGGTC CATCAATTAA AAACCTGGTG AACATAAAAT TAACCAAGCA
--> 3158 ATCAATTTAG 690
3159
3160 ACAAGCAACA TAATCATATA TATTTTCTT ACATTTGTAT GTACCTGAAT
--> 3161 ATTTATATTT 750
3162
3163 ATGTTTATAT GTTCTCACTA TATTTTCACT TTTGTATTG AAAATTTTGA GGA AAA
3164 806
3165
3166 Gly Lys
--> 3167 ATT TTG GAT AGA AAA TTG ATG GGA GAA GAT CTC TAC TGG GCA ATA AAC 854
3168 Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn
3169 195 200 205
3170
3171 GGA GGA GGA GGA GGG AGC TAC GGC GTC GTA TTG GCC TAC AAA ATA AAC ←
3172 902
3173 Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn
3174 210 215 220
3175
--> 3176 CTT GTT GAA GTC CCA GAA AAC GTC ACC GTT TTC AGA ATC TCC CGG ACG 950
3177 Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr
3178 225 230 235 240
3179
3180 TTA GAA CAA AAT GCG ACG GAT ATC ATT CAC CGG TGG CAA CAA GTT GCA 998
3181 Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala
3182 245 250 255
3183
3184 CCG AAG CTT CCC GAC GAG CTT TTC ATA AGA ACA GTC ATT GAC GTA GTA 1046
3185 Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val
3186 260 265 270
3187
3188
3189 AAC GGC ACT GTT TCA TCT CAA AAG ACC GTC AGG ACA ACA TTC ATA GCA ←
3190 1094
3191 Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala
3192 275 280 285
3193
--> 3194 ATG TTT CTA GGA GAC ACG ACA ACT CTA CTG TCG ATA TTA AAC CGG AGA 1142
3195 Met Phe Leu Gly Asp Thr Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg
3196 290 295 300
3197
3198 TTC CCA GAA TTG GGT TTG GTC CGG TCT GAC TGT ACC GAA ACA AGC TGG 1190
3199 Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp
3200 305 310 315 320
3201
3202 ATC CAA TCT GTG CTA TTC TGG ACA AAT ATC CAA GTT GGT TCG TCG GAG 1238
3203 Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu
3204 325 330 335
3205
3206 ACA CTT CTA CTC CAA AGG AAT CAA CCC GTG AAC TAC CTC AAG AGG AAA ←
3207 1286
3208 Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys
3209 340 345 350

```

*sane*

←

←

←

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/258,031ADATE: 09/03/1999  
TIME: 14:51:48

INPUT SET: S33211.raw

3210  
--> 3211 TCA GAT TAC GTA CGT GAA CCG ATT TCA AGA ACC GGT TTA GAG TCA ATT 1334  
3212 Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile  
3213 355 360 365  
3214  
3215 TGG AAG AAA ATG ATC GAG CTT GAA ATT CCG ACA ATG GCT TTC AAT CCA  
3216 1382  
3217 Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro  
3218 370 375 380  
3219  
3220 TAC GGT GGT GAG ATG GGG AGG ATA TCA TTA CGG GTG ACT CCG TTC CCA  
3221 1430  
3222 Tyr Gly Gly Glu Met Gly Arg Ile Ser Leu Arg Val Thr Pro Phe Pro  
3223 385 390 395 400  
3224  
3225 TAC AGA GCC GGT AAT CTC TGG AAG ATT CAG TAC GGT GCG AAT TGG AGA  
3226 1478  
3227 Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg  
3228 405 410 415  
3229  
3230 GAT GAG ACT TTA ACC GAC CGG TAC ATG GAA TTG ACG AGG AAG TTG TAC  
3231 1526  
3232 Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr  
3233 420 425 430  
3234  
--> 3235 CAA TTC ATG ACA CCA TTT GTT TCC AAG AAT CCG AGA CAA TCG TTT TTC 1574  
3236 Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe  
3237 435 440 445  
3238  
3239 AAT AAC CGT GAT GTT GAT TTG GGT ATT AAT TCT CAT AAT GGT AAA ATC 1622  
3240 Asn Asn Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile  
3241 450 455 460  
3242  
3243 AGT AGT TAT GTG GAA GGT AAA CGT TAC GGG AAG AAG TAT TTC GCA GGT  
3244 1670  
3245 Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly  
3246 465 470 475 480  
3247  
3248 AAT TTC GAG AGA TTG GTG AAG ATT AAG ACG AGA GTT GAT AGT GGT AAT  
3249 1718  
3250 Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn  
3251 485 490 495  
3252  
3253  
3254  
--> 3255 TTC TTT AGG AAC GAA CAC AGT ATT CCT GTG TTA CCA TAA 1757  
3256 Phe Phe Arg Asn Glu His Ser Ile Pro Val Leu Pro  
3257 500 505  
3258  
3259

---

3368 (2) INFORMATION FOR SEQ ID NO: 72:  
3369

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:48

INPUT SET: S33211.raw

```

--> 3370      (i) SEQUENCE CHARACTERISTICS:
3371          (A) LENGTH: 1527 base pairs
3372          (B) TYPE: nucleic acid
3373          (C) STRANDEDNESS: double
3374          (D) TOPOLOGY: linear
3375
3376      (ii) MOLECULE TYPE: cDNA to mRNA
3377
3378      (iii) HYPOTHETICAL: NO
3379
3380      (iii) ANTI-SENSE: NO
3381
3382      (vi) ORIGINAL SOURCE:
3383          (A) ORGANISM: Arabidopsis thaliana
3384          (B) STRAIN: Colombia
3385
3386      (ix) FEATURE:
3387          (A) NAME/KEY: CDS
3388          (B) LOCATION: 1..1524
3389
3390      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:
3391
3392      ACT TCC CGT AGA AAC TCG GAG ACT TTC ACA CAA TGC CTA ACC TCA AAC      48
3393      Thr Ser Arg Arg Asn Ser Glu Thr Phe Thr Gln Cys Leu Thr Ser Asn
3394      1          5          10          15
3395
3396      TCC GAC CCC AAA CAT CCC ATC TCC CCC GCT ATC TTC TTC TCC GGA AAT      96
3397      Ser Asp Pro Lys His Pro Ile Ser Pro Ala Ile Phe Phe Ser Gly Asn
3398      20          25          30
3399
3400      GGC TCC TAC TCC TCC GTA TTA CAA GCC AAC ATC CGT AAC CTC CGC TTC      144
3401      Gly Ser Tyr Ser Ser Val Leu Gln Ala Asn Ile Arg Asn Leu Arg Phe
3402      35          40          45
3403
3404      AAC ACC ACC TCA ACT CCG AAA CCC TTC CTC ATA ATC GCC GCA ACA CAT      192
3405      Asn Thr Thr Ser Thr Pro Lys Pro Phe Leu Ile Ile Ala Ala Thr His
3406      50          55          60
3407
3408      GAA TCC CAT GTG CAA GCC GCG ATT ACT TGC GGG AAA CGC CAC AAC CTT      240
3409      Glu Ser His Val Gln Ala Ala Ile Thr Cys Gly Lys Arg His Asn Leu
3410      65          70          75          80
3411
3412      CAG ATG AAA ATC AGA AGT GGA GGC CAC GAC TAC GAT GGC TTG TCA TAC
3413      288
3414      Gln Met Lys Ile Arg Ser Gly Gly His Asp Tyr Asp Gly Leu Ser Tyr
3415      85          90          95
3416
--> 3417      GTT ACA TAC TCT GGC AAA CCG TTC TTC GTC CTC GAC ATG TTT AAC CTC      336
3418      Val Thr Tyr Ser Gly Lys Pro Phe Phe Val Leu Asp Met Phe Asn Leu
3419      100          105          110
3420
3421      CGT TCG GTG GAT GTC GAC GTG GCA AGT AAG ACC GCG TGG GTC CAA ACC      384
3422      Arg Ser Val Asp Val Asp Val Ala Ser Lys Thr Ala Trp Val Gln Thr

```

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

 DATE: 09/03/1999  
 TIME: 14:51:49

INPUT SET: S33211.raw

	115	120	125	
3423				
3424				
3425	GGT GCC ATA CTC GGA GAA GTT TAT TAC TAT ATA TGG GAG AAG AGC AAA			432
3426	Gly Ala Ile Leu Gly Glu Val Tyr Tyr Tyr Ile Trp Glu Lys Ser Lys			
3427	130	135	140	
3428				
3429	ACC CTA GCT TAT CCC GCC GGA ATT TGT CCC ACG GTT GGT GTC GGT GGC			480
3430	Thr Leu Ala Tyr Pro Ala Gly Ile Cys Pro Thr Val Gly Val Gly Gly			
3431	145	150	155	160
3432				
3433	CAT ATC AGT GGT GGA GGT TAC GGT AAC ATG ATG AGA AAA TAC GGT CTC			528
3434	His Ile Ser Gly Gly Gly Tyr Gly Asn Met Met Arg Lys Tyr Gly Leu			
3435	165	170	175	
3436				
3437	ACC GTA GAT AAT ACC ATC GAT GCA AGA ATG GTC GAC GTA AAT GGA AAA			
3438	576			
3439	Thr Val Asp Asn Thr Ile Asp Ala Arg Met Val Asp Val Asn Gly Lys			
3440	180	185	190	
3441				
--> 3442	ATT TTG GAT AGA AAA TTG ATG GGA GAA GAT CTC TAC TGG GCA ATA AAC			624
3443	Ile Leu Asp Arg Lys Leu Met Gly Glu Asp Leu Tyr Trp Ala Ile Asn			
3444	195	200	205	
3445				
3446	GGA GGA GGA GGA GGG AGC TAC GGC GTC GTA TTG GCC TAC AAA ATA AAC			
3447	672			
3448	Gly Gly Gly Gly Gly Ser Tyr Gly Val Val Leu Ala Tyr Lys Ile Asn			
3449	210	215	220	
3450				
--> 3451	CTT GTT GAA GTC CCA GAA AAC GTC ACC GTT TTC AGA ATC TCC CGG ACG			720
3452	Leu Val Glu Val Pro Glu Asn Val Thr Val Phe Arg Ile Ser Arg Thr			
3453	225	230	235	240
3454				
3455	TTA GAA CAA AAT GCG ACG GAT ATC ATT CAC CGG TGG CAA CAA GTT GCA			768
3456	Leu Glu Gln Asn Ala Thr Asp Ile Ile His Arg Trp Gln Gln Val Ala			
3457	245	250	255	
3458				
3459	CCG AAG CTT CCC GAC GAG CTT TTC ATA AGA ACA GTC ATT GAC GTA GTA			816
3460	Pro Lys Leu Pro Asp Glu Leu Phe Ile Arg Thr Val Ile Asp Val Val			
3461	260	265	270	
3462				
3463	AAC GGC ACT GTT TCA TCT CAA AAG ACC GTC AGG ACA ACA TTC ATA GCA			864
3464	Asn Gly Thr Val Ser Ser Gln Lys Thr Val Arg Thr Thr Phe Ile Ala			
3465	275	280	285	
3466				
3467	ATG TTT CTA GGA GAC ACG ACA ACT CTA CTG TCG ATA TTA AAC CGG AGA			912
3468	Met Phe Leu Gly Asp Thr Thr Thr Leu Leu Ser Ile Leu Asn Arg Arg			
3469	290	295	300	
3470				
3471	TTC CCA GAA TTG GGT TTG GTC CGG TCT GAC TGT ACC GAA ACA AGC TGG			960
3472	Phe Pro Glu Leu Gly Leu Val Arg Ser Asp Cys Thr Glu Thr Ser Trp			
3473	305	310	315	320
3474				
3475	ATC CAA TCT GTG CTA TTC TGG ACA AAT ATC CAA GTT GGT TCG TCG GAG			1008

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:49

INPUT SET: S33211.raw

```

3476 Ile Gln Ser Val Leu Phe Trp Thr Asn Ile Gln Val Gly Ser Ser Glu
3477             325             330             335
3478
3479 ACA CTT CTA CTC CAA AGG AAT CAA CCC GTG AAC TAC CTC AAG AGG AAA
3480 1056
3481 Thr Leu Leu Leu Gln Arg Asn Gln Pro Val Asn Tyr Leu Lys Arg Lys
3482             340             345             350
3483
--> 3484 TCA GAT TAC GTA CGT GAA CCG ATT TCA AGA ACC GGT TTA GAG TCA ATT 1104
3485 Ser Asp Tyr Val Arg Glu Pro Ile Ser Arg Thr Gly Leu Glu Ser Ile
3486             355             360             365
3487
3488 TGG AAG AAA ATG ATC GAG CTT GAA ATT CCG ACA ATG GCT TTC AAT CCA
3489 1152
3490 Trp Lys Lys Met Ile Glu Leu Glu Ile Pro Thr Met Ala Phe Asn Pro
3491             370             375             380
3492
--> 3493 TAC GGT GGT GAG ATG GGG AGG ATA TCA TCT ACG GTG ACT CCG TTC CCA 1200
3494 Tyr Gly Gly Glu Met Gly Arg Ile Ser Ser Thr Val Thr Pro Phe Pro
3495 385             390             395             400
3496
3497 TAC AGA GCC GGT AAT CTC TGG AAG ATT CAG TAC GGT GCG AAT TGG AGA
3498 1248
3499 Tyr Arg Ala Gly Asn Leu Trp Lys Ile Gln Tyr Gly Ala Asn Trp Arg
3500             405             410             415
3501
3502 GAT GAG ACT TTA ACC GAC CGG TAC ATG GAA TTG ACG AGG AAG TTG TAC
3503 1296
3504 Asp Glu Thr Leu Thr Asp Arg Tyr Met Glu Leu Thr Arg Lys Leu Tyr
3505             420             425             430
3506
--> 3507 CAA TTC ATG ACA CCA TTT GTT TCC AAG AAT CCG AGA CAA TCG TTT TTC 1344
3508 Gln Phe Met Thr Pro Phe Val Ser Lys Asn Pro Arg Gln Ser Phe Phe
3509             435             440             445
3510
3511 AAT TAC CGT GAT GTT GAT TTG GGT ATT AAT TCT CAT AAT GGT AAA ATC 1392
3512 Asn Tyr Arg Asp Val Asp Leu Gly Ile Asn Ser His Asn Gly Lys Ile
3513             450             455             460
3514
3515 AGT AGT TAT GTG GAA GGT AAA CGT TAC GGG AAG AAG TAT TTC GCA GGT
3516 1440
3517 Ser Ser Tyr Val Glu Gly Lys Arg Tyr Gly Lys Lys Tyr Phe Ala Gly
3518 465             470             475             480
3519
3520 AAT TTC GAG AGA TTG GTG AAG ATT AAG ACG AGA GTT GAT AGT GGT AAT
3521 1488
3522 Asn Phe Glu Arg Leu Val Lys Ile Lys Thr Arg Val Asp Ser Gly Asn
3523             485             490             495
3524
--> 3525 TTC TTT AGG AAC GAA CAG AGT ATT CCT GTG TTA CCA TAA 1527
3526 Phe Phe Arg Asn Glu Gln Ser Ile Pro Val Leu Pro
3527             500             505
3528

```

*same*

*same*

*same*

*same*

INPUT SET: S33211.raw

3529

```

3639 (2) INFORMATION FOR SEQ ID NO: 74:
3640
3641 (i) SEQUENCE CHARACTERISTICS:
--> 3642 (A) LENGTH: 1530 base pairs
3643 (B) TYPE: nucleic acid
3644 (C) STRANDEDNESS: double
3645 (D) TOPOLOGY: linear
3646
3647 (ii) MOLECULE TYPE: cDNA to mRNA
3648
3649 (iii) HYPOTHETICAL: NO
3650
3651 (iii) ANTI-SENSE: NO
3652
3653 (vi) ORIGINAL SOURCE:
3654 (A) ORGANISM: Arabidopsis thaliana
3655 (B) STRAIN: Colombia
3656
3657 (ix) FEATURE:
3658 (A) NAME/KEY: CDS
3659 (B) LOCATION: 1..1527
3660
3661 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:
3662
3663 TCC ATC CAA GAT CAA TTC ATA AAC TGT GTC AAA AGA AAC ACA CAT GTT 48
3664 Ser Ile Gln Asp Gln Phe Ile Asn Cys Val Lys Arg Asn Thr His Val
3665 1 5 10 15
3666
3667 TCT TTT CCA CTC GAG AAA ACG TTA TTC ACC CCT GCG AAA AAC GTC TCT 96
3668 Ser Phe Pro Leu Glu Lys Thr Leu Phe Thr Pro Ala Lys Asn Val Ser
3669 20 25 30
3670
3671
3672 TTG TTC AAC CAA GTC CTT GAA TCG ACG GCT CAA AAT CTC CAG TTC TTG 144
3673 Leu Phe Asn Gln Val Leu Glu Ser Thr Ala Gln Asn Leu Gln Phe Leu
3674 35 40 45
3675
3676 GCA AAA TCC ATG CCT AAA CCG GGA TTC ATA TTC AGA CCG ATT CAC CAG 192
3677 Ala Lys Ser Met Pro Lys Pro Gly Phe Ile Phe Arg Pro Ile His Gln
3678 50 55 60
3679
3680 TCT CAA GTC CAA GCT TCC ATC ATT TGT TCA AAG AAA CTC GGA ATT CAT 240
3681 Ser Gln Val Gln Ala Ser Ile Ile Cys Ser Lys Lys Leu Gly Ile His
3682 65 70 75 80
3683
3684 TTT CGT GTT AGA AGT GGC GGT CAC GAT TTC GAG GCC TTG TCT TAT GTT 288
3685 Phe Arg Val Arg Ser Gly Gly His Asp Phe Glu Ala Leu Ser Tyr Val
3686 85 90 95
3687
3688 TCA CGG ATT GAA AAA CCG TTT ATA TTA CTC GAC CTG TCA AAA TTG AAA 336
3689 Ser Arg Ile Glu Lys Pro Phe Ile Leu Leu Asp Leu Ser Lys Leu Lys

```

pp 46-47

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:50

INPUT SET: S33211.raw

	100	105	110	
3690				
3691				
3692	CAA ATC AAT GTT GAT ATT GAA TCC AAT AGT GCT TGG GTT CAA CCT GGT			384
3693	Gln Ile Asn Val Asp Ile Glu Ser Asn Ser Ala Trp Val Gln Pro Gly			
3694	115	120	125	
3695				
3696	GCT ACG CTT GGT GAG CTT TAC TAC AGA ATT GCA GAG AAG AGC AAG ATC			432
3697	Ala Thr Leu Gly Glu Leu Tyr Tyr Arg Ile Ala Glu Lys Ser Lys Ile			
3698	130	135	140	
3699				
3700	CAT GGA TTT CCC GCG GGT TTG TGC ACA AGT GTA GGC ATA GGT GGG TAT			480
3701	His Gly Phe Pro Ala Gly Leu Cys Thr Ser Val Gly Ile Gly Gly Tyr			
3702	145	150	155	160
3703				
3704	ATG ACA GGC GGT GGA TAC GGT ACC TTG ATG AGG AAG TAT GGT CTT GCG			528
3705	Met Thr Gly Gly Gly Tyr Gly Thr Leu Met Arg Lys Tyr Gly Leu Ala			
3706	165	170	175	
3707				
3708	GGA GAT AAT GTT CTA GAC GTA AAG ATG GTT GAT GCA AAT GGT AAA TTA			576
3709	Gly Asp Asn Val Leu Asp Val Lys Met Val Asp Ala Asn Gly Lys Leu			
3710	180	185	190	
3711				
3712	CTC GAC AGA GCC GCG ATG GGT GAG GAC CTA TTT TGG GCG ATT AGA GGA			
3713	624			
3714	Leu Asp Arg Ala Ala Met Gly Glu Asp Leu Phe Trp Ala Ile Arg Gly			
3715	195	200	205	
3716				
--> 3717	GGC GGT GGA GCG AGT TTC GGG ATA GTT CTA GCA TGG AAG ATC AAG CTT			672
3718	Gly Gly Gly Ala Ser Phe Gly Ile Val Leu Ala Trp Lys Ile Lys Leu			
3719	210	215	220	
3720				
3721	GTT CCT GTT CCT AAG ACT GTT ACC GTC TTC ACT GTC ACC AAA ACG TTA			720
3722	Val Pro Val Pro Lys Thr Val Thr Val Phe Thr Val Thr Lys Thr Leu			
3723	225	230	235	240
3724				
3725	GAA CAA GAC GCA AGA TTG AAG ACT ATT TCT AAG TGG CAA CAA ATT TCA			768
3726	Glu Gln Asp Ala Arg Leu Lys Thr Ile Ser Lys Trp Gln Gln Ile Ser			
3727	245	250	255	
3728				
3729				
3730				
3731	TCC AAG ATT ATT GAA GAG ATA CAC ATC CGA GTG GTA CTC AGA GCA GCT			816
3732	Ser Lys Ile Ile Glu Glu Ile His Ile Arg Val Val Leu Arg Ala Ala			
3733	260	265	270	
3734				
3735	GGA AAT GAT GGA AAC AAG ACT GTG ACA ATG ACC TAC CTA GGT CAG TTT			864
3736	Gly Asn Asp Gly Asn Lys Thr Val Thr Met Thr Tyr Leu Gly Gln Phe			
3737	275	280	285	
3738				
3739	CTT GGC GAG AAA GGC ACC TTG CTG AAG GTT ATG GAG AAG GCT TTT CCA			912
3740	Leu Gly Glu Lys Gly Thr Leu Leu Lys Val Met Glu Lys Ala Phe Pro			
3741	290	295	300	
3742				

*same*

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/258,031A

DATE: 09/03/1999  
TIME: 14:51:50

INPUT SET: S33211.raw

3743	GAA CTA GGG TTA ACT CAA AAG GAT TGT ACT GAA ATG AGC TGG ATT GAA	960
3744	Glu Leu Gly Leu Thr Gln Lys Asp Cys Thr Glu Met Ser Trp Ile Glu	
3745	305 310 315 320	
3746		
3747	GCC GCC CTT TTC CAT GGT GGA TTT CCA ACA GGT TCT CCT ATT GAA ATT	1008
3748	Ala Ala Leu Phe His Gly Gly Phe Pro Thr Gly Ser Pro Ile Glu Ile	
3749	325 330 335	
3750		
3751	TTG CTT CAG CTC AAG TCG CCT CTA GGA AAA GAT TAC TTC AAA GCA ACG	1056
3752	Leu Leu Gln Leu Lys Ser Pro Leu Gly Lys Asp Tyr Phe Lys Ala Thr	
3753	340 345 350	
3754		
3755	TCG GAT TTC GTT AAA GAA CCT ATT CCT GTG ATA GGC TTC AAA GGA ATA	1104
3756	Ser Asp Phe Val Lys Glu Pro Ile Pro Val Ile Gly Phe Lys Gly Ile	
3757	355 360 365	
3758		
3759	TTC AAA AGA TTG ATT GAA GGA AAC ACA ACA TTT CTG AAC TGG ACT CCT	1152
3760	Phe Lys Arg Leu Ile Glu Gly Asn Thr Thr Phe Leu Asn Trp Thr Pro	
3761	370 375 380	
3762		
3763	TAC GGT GGT ATG ATG TCG AAA ATC CCT GAA TCT GCG ATC CCA TTT CCG	1200
3764	Tyr Gly Gly Met Met Ser Lys Ile Pro Glu Ser Ala Ile Pro Phe Pro	
3765	385 390 395 400	
3766		
3767	CAT AGA AAC GGA ACC CTC TTC AAG ATT CTC TAT TAC GCG AAC TGG CTA	1248
3768	His Arg Asn Gly Thr Leu Phe Lys Ile Leu Tyr Tyr Ala Asn Trp Leu	
3769	405 410 415	
3770		
3771	GAG AAT GAC AAG ACA TCG AGT AGA AAA ATC AAC TGG ATC AAA GAG ATA	
3772	1296	
3773	Glu Asn Asp Lys Thr Ser Ser Arg Lys Ile Asn Trp Ile Lys Glu Ile	
3774	420 425 430	
3775		
--> 3776	TAC AAT TAC ATG GCG CCT TAT GTC TCA AGC AAT CCA AGA CAA GCA TAT	1344
3777	Tyr Asn Tyr Met Ala Pro Tyr Val Ser Ser Asn Pro Arg Gln Ala Tyr	
3778	435 440 445	
3779		
3780	GTG AAC TAC AGA GAT CTA GAC TTC GGA CAG AAC AAG AAC AAC GCA AAG	
3781	1392	
3782	Val Asn Tyr Arg Asp Leu Asp Phe Gly Gln Asn Lys Asn Asn Ala Lys	
3783	450 455 460	
--> 3784	GTT AAC TTC ATT GAA GCT AAA ATC TGG GGA CCT AAG TAC TTC AAA GGC	1440
3785	Val Asn Phe Ile Glu Ala Lys Ile Trp Gly Pro Lys Tyr Phe Lys Gly	
3786	465 470 475 480	
3787		
3788	AAT TTT GAC AGA TTG GTG AAG ATT AAA ACC AAG GTT GAT CCA GAG AAC	
3789	1488	
3790	Asn Phe Asp Arg Leu Val Lys Ile Lys Thr Lys Val Asp Pro Glu Asn	
3791	485 490 495	
3792		
--> 3793	TTC TTC AGG CAC GAG CAG AGT ATC CCA CCT ATG CCC TAC TAG	1530
3794	Phe Phe Arg His Glu Gln Ser Ile Pro Pro Met Pro Tyr	
3795	500 505	